

## PCT COOPERATION TREATY

From the INTERNATIONAL BUREAU

PCT

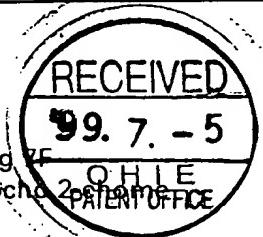
**NOTIFICATION CONCERNING  
SUBMISSION OR TRANSMITTAL  
OF PRIORITY DOCUMENT**

(PCT Administrative Instructions, Section 411)

Date of mailing (day/month/year)  
23 June 1999 (23.06.99)

To:

OHIE, Kunihisa  
Ohie Patent Office  
Horiguchi No. 2 Building 2F  
2-6, Nihonbashi-Ningyocho  
Chuo-ku  
Tokyo 103-0013  
JAPON



Applicant's or agent's file reference  
ONF-2970PCT

**IMPORTANT NOTIFICATION**

International application No.  
PCT/JP99/02284

International filing date (day/month/year)  
28 April 1999 (28.04.99)

International publication date (day/month/year)  
Not yet published

Priority date (day/month/year)  
28 April 1998 (28.04.98)

Applicant

ONO PHARMACEUTICAL CO., LTD. et al

1. The applicant is hereby notified of the date of receipt (except where the letters "NR" appear in the right-hand column) by the International Bureau of the priority document(s) relating to the earlier application(s) indicated below. Unless otherwise indicated by an asterisk appearing next to a date of receipt, or by the letters "NR", in the right-hand column, the priority document concerned was submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b).
2. This updates and replaces any previously issued notification concerning submission or transmittal of priority documents.
3. An asterisk(\*) appearing next to a date of receipt, in the right-hand column, denotes a priority document submitted or transmitted to the International Bureau but not in compliance with Rule 17.1(a) or (b). In such a case, the attention of the applicant is directed to Rule 17.1(c) which provides that no designated Office may disregard the priority claim concerned before giving the applicant an opportunity, upon entry into the national phase, to furnish the priority document within a time limit which is reasonable under the circumstances.
4. The letters "NR" appearing in the right-hand column denote a priority document which was not received by the International Bureau or which the applicant did not request the receiving Office to prepare and transmit to the International Bureau, as provided by Rule 17.1(a) or (b), respectively. In such a case, the attention of the applicant is directed to Rule 17.1(c) which provides that no designated Office may disregard the priority claim concerned before giving the applicant an opportunity, upon entry into the national phase, to furnish the priority document within a time limit which is reasonable under the circumstances.

<u>Priority date</u>	<u>Priority application No.</u>	<u>Country or regional Office or PCT receiving Office</u>	<u>Date of receipt of priority document</u>
28 April 1998 (28.04.98)	10/119731	JP	22 June 1999 (22.06.99)

The International Bureau of WIPO  
34, chemin des Colombettes  
1211 Geneva 20, Switzerland

Facsimile No. (41-22) 740.14.35

Authorized officer

Juan Cruz

Telephone No. (41-22) 338.83.38

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

D E C L A R A T I O N

I, Tetsuya Kato, declare:

1. That I am a Japanese subject, residing at c/o Minase Research Institute, Ono Pharmaceutical Co., Ltd., 1-1, Sakurai 3-chome, Shimamoto-cho, Mishima-gun, OSAKA, JAPAN.
2. That I am well acquainted with the Japanese and English languages.
3. That the attached is true translation into the English language of the accompanying certified copy of the Application for Patent filed in Japan on April 28th, 1998, under the number Heisei 10-119731.
4. That all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements are made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the Patent Application the United States of America of any patent issuing thereon.

Dated this 16th day of October, 2000.

  
Tetsuya Kato

Translation:

Patent Office  
Japanese Government

This is to certify that the annexed is a true copy of the following application as filed this Office.

Date of Application: April 28th, 1998  
Application number : Patent Application No. Heisei 10-119731  
Applicant(s) : Ono Pharmaceutical Co., Ltd.

Dated this May 28th, 1999

Takeshi Isayama

-----  
Commissioner,

Patent Office

Certificate No. Patent Heisei  
11-3034414

Document Name: Application for Patent

Reference No.: GEJP-52

Submitted date: 27th April, 1998

Direction: Commissioner, Patent Office

IPC: C12N 15/00

C07K 7/00

Title of the Invention: A novel polypeptide, a method of producing it, a cDNA encoding it, a vector containing it, a host cell transformed with the vector, an antibody of the peptide, a pharmaceutical composition containing the polypeptide or the antibody, a screening method with using the polypeptide

Number of claims: 13

Inventor:

Address: 19-4, Ohsagi-cho, Iwakura, Sakyo-ku, Kyoto.

Name: Tasuku Honjo

Inventor:

Address: 93, Higashiohno-cho, Koyama, Kita-ku, Kyoto

Name: Kei Tashiro

Inventor:

Address: 7665, Palmira Dr., #5324, San Diego, CA, U.S.A.

Name: Tomoyuki Nakamura

Applicant:

ID NO.: 000185983

Zip code: 541

Address: 2-1-5, Doshomachi, Chuo-ku, Osaka

Name: Ono Pharmaceutical Co., Ltd.

Representative: Toshio Ueno

Charge:

Method for payment: prepaid

No. of Ledger: 029595

Sum of prepaid: 21,000

List of Attached document:

Item: Specification 1

Item: Figures 2

Item: Abstract 1

Proof: Yes

**Document Name:** Specification

**Title of the Invention:**

A novel polypeptide, a method for preparation of it, a cDNA encoding the polypeptide

**Claims**

1. Substantially purified form of the polypeptide that comprising the amino-acid sequence shown in SEQ ID NO. 1, 4, 6, 9, 11 or 14, homologue thereof, fragment thereof or homologue of the fragment.

2. A polypeptide according to claim 1 that comprising the amino-acid sequence shown in SEQ ID NO. 1, 4, 6, 9, 11 or 14.

3. A cDNA encoding the polypeptide according to claim 1.

4. A cDNA according to claim 3 that comprising the nucleotide sequence shown in SEQ ID NO. 2, 5, 7, 10, 12 or 15 or a fragment cDNA selectively hybridized to the cDNA.

5. A cDNA according to claim 3 that comprising the nucleotide sequence shown in SEQ ID NO. 3, 8 or 13 or a fragment cDNA selectively hybridized to the cDNA.

6. A replication or expression vector carrying the cDNA according to claim 3 to 5.

7. A host cell transformed with the replication or expression vector according to claim 6.

8. A method for producing the polypeptide according to claim 1 or 2 which comprises culturing a host cell according to claim 7 under a condition effective to express the polypeptide according to claim 1 or 2.

9. A monoclonal or polyclonal antibody against the polypeptide according to claim 1 or 2.

10. A pharmaceutical composition containing the polypeptide

according to claim 1 or 2 or the antibody according to claim 9, in association with pharmaceutically acceptable diluent and/or carrier.

11. A pharmaceutical composition for the treatment of abnormal growth of smooth muscle cell, containing a polypeptide according to claim 1 or 2, in association with a pharmaceutically acceptable diluent and/or carrier.

12. A pharmaceutical composition for the treatment of arteriosclerosis, restenosis after PTCA or myosarcoma, containing the polypeptide according to claim 1 or 2, in association with a pharmaceutically acceptable diluent and/or carrier.

13. A screening method for an antagonist or agonist of the polypeptide according to claim 1 or 2 with using the said polypeptide.

## **Detailed Description of the Invention**

### **Field of the Invention**

The present invention provides a novel polypeptide, a method for preparation of them, a cDNA encoding the polypeptide, a vector containing it, a host cell transformed with the vector, an antibody of the peptide, a pharmaceutical composition containing the polypeptide or the antibody, a screening method with using the polypeptide.

### **Problems to be Solved**

The present inventors investigated to find novel factors (polypeptides) which are useful for study or for the treatment or diagnosis of diseases induced by abnormal proliferation of smooth muscle. Especially, we had aimed secreted proteins and membrane proteins which have signal sequences for secretion.

### **Background of the Invention**

In modern medical research, cardiovascular biology is a field that attracts considerable attention because cardiovascular disease is the leading cause of mortality. Cardiovascular research has revealed important facts about neointimal formation and arterial remodeling, both of which are thought to contribute to plaque formation in atherosclerosis and blood vessel narrowing. For example, there are three aspects of the cellular process in hypercholesterolaemia induced blood vessel damage in animal models that mimic human development of arteriosclerotic coronary disease. The three elements that form lesions on the artery wall are: a) proliferation of smooth muscle cells, macrophages and lymphocytes, b) formation of connective tissues (mainly elastic fiber proteins, collagen and proteoglycans made by smooth muscle cells in a process similar to scar formation), and c) the accumulation

of lipid and cholesterol in the newly formed connective tissue matrices. The exact sequence of the three damaging elements are debatable, but it is clear that the abnormal dedifferentiation, redifferentiation and growth of smooth muscle cells contribute structurally to vessel damage. Moreover, another significant pathological process that involves abnormal smooth muscle cell growth is restenosis after Percutaneous transluminal coronary angioplasty (PTCA).

The present inventors made reasonable efforts, by isolation of the molecules related to participation of smooth muscle cells in angiogenesis, for the aim to utilize them for regulation of abnormal proliferation of smooth muscle cells such like described above.

In order to obtain a certain polypeptide or cDNA coding for the same, there has been generally employed a method composed of detecting the aimed biological activity in a tissue or a cell culture medium, then identifying a polypeptide as substance of the activity through the isolation and purification and isolating a gene encoding the polypeptide or expression-cloning method to isolate a gene by access of the biological activity of the polypeptide encoded by it.

Because in many cases, however, physiologically active polypeptides have various biological activities, when taking the method to approaches based on a certain activity to isolate a gene, it has increasingly been happened that the gene is turned out to be identical to a known gene which has another activity after spending much efforts to isolate it. And because, in many cases, biological factors are produced only in a very slight amount or only in a specific condition, it is often made difficult to isolate and purify a factor and detect its biological activity.

#### **Related Arts**

Recent rapid developments in techniques for constructing cDNAs and sequencing techniques have made it possible to quickly sequence a large amount of cDNAs. By utilizing these techniques, a process, which comprises constructing cDNAs at random, identifying the nucleotide sequences thereof, expressing novel polypeptides encoded by them, is now in progress. Although this process is advantageous in that a gene can be cloned and information regarding its nucleotide sequence can be obtained without any biochemical or genetic analysis, the target gene can be discovered thereby only accidentally in many cases.

#### **Means for solving the problems**

The present inventors have studied cloning method of genes coding proliferation and/or differentiation factors functioning in hematopoietic systems and immune systems. Focusing their attention on the fact that most of the secretory proteins such as proliferation and/or differentiation factors (for example various cytokines) and membrane proteins such as receptors thereof (hereafter these proteins will be referred to generally as secretory proteins and the like) have sequences called signal peptides in the N-termini, the inventors conducted extensive studies on a process for efficiently and selectively cloning a gene coding for a signal peptide. Finally, we have successfully invented a screening method for cDNAs having sequence encoding signal peptides, we called the method as signal sequence trap (SST) (Japanese Patent Application No. 6-13951).

We also developed yeast SST method on the same concept. By the method using yeast, genes including sequence encoding signal peptide can be identified more easily and effectively (USP No. 5,536,637).

By using the present method, the present inventors identified novel secreted protein produced by mouse embryonic heart and human kidney and a

cDNA fragments encoding them, and by using the sequence information of the cDNA fragments they isolated each full-length cDNA from mouse embryonic heart and human kidney. And they discovered that the polypeptides had functions to suppress smooth muscle cells.

The present cDNA sequence was identified as a clone mouse A55 and isolated from cDNA library derived from mouse embryonic heart based on genetic information obtained by using the Yeast SST method described above. The clone, mouse A55 is a full-length cDNA encoding a secreted polypeptide (which is called mouse A55 polypeptide here).

The present cDNA sequence was named as a clone human A55 and isolated from cDNA library derived from human brain based on genetic information obtained from human kidney by using the Yeast SST method described above. The clone, human A55 is a full-length cDNA encoding a secreted polypeptide (which is called human A55 polypeptide here).

There was no DNA sequence which is identical to that of mouse and human A55 of the present invention, when DNA sequence of mouse and human A55 were compared with data base by BLASTN and FASTA. And there was no polypeptides which is identical to that of mouse and human A55 of the present invention, when amino acid sequence of mouse and human A55 was compared with data base by BLASTX, BLASTP and FASTA. So the polypeptides of the present invention are considered to be novel.

The inventors discovered that the polypeptides had functions to suppress smooth muscle cells. Accordingly, the polypeptides may be useful for treatment of diseases related to abnormal proliferation of smooth muscle cells, for example, arteriosclerotic coronary disease, neointimal formation which results in restenosis after percutaneous transluminal coronary angioplasty and myosarcoma.

## **Constitution of the Invention**

The present invention provides:

- 1) a polypeptide having an amino acid sequence shown in SEQ ID NO. 1, 4, 6, 9, 11 or 14,
- 2) a cDNA encoding the polypeptide described above (1),
- 3) a cDNA having an nucleotide sequence shown in SEQ ID NO. 2, 5, 7, 10, 12 or 15,
- 4) a cDNA that consists of an nucleotide sequence shown in SEQ ID NO. 3, 8 or 13.

## **Detailed Discription**

The present invention is concerned with a polypeptide that consists of the amino acid sequence shown in SEQ ID NO. 1, 4, 6, 9, 11 or 14 in substantially purified form, a homologue thereof, a fragment of the sequence and a homologue of the fragment.

Further, the present invention is concerned with a cDNA encoding the above peptides. More particularly the present invention is provided cDNA having the nucleotide sequence shown in SEQ ID NO. 2, 5, 7, 10, 12 or 15, and cDNA containing a fragment which is selectively hybridizing to the cDNA that consists of nucleotide sequence shown in SEQ ID NO. 2, 5, 7, 10, 12 or 15. Complementary sequence of the above nucleotide sequence is also included in cDNA selectively hybridized. Hybridization are performed in an stringent condition.

A polypeptide comprising amino acid sequence shown in SEQ ID NO. 1, 4, 6, 9, 11 or 14 in substantially purified form will generally comprise the polypeptide in a preparation in which more than 90%, e.g. 95%, 98% or 99% of the polypeptide in the preparation is that of the SEQ ID NO. 1, 4, 6, 9, 11 or 14.

A homologue of polypeptide comprising amino acid sequence shown in SEQ ID NO. 1, 4, 6, 9, 11 or 14 will be generally at least 70%, preferably at least 80 or 90% and more preferably at least 95% homologous to the polypeptide of SEQ ID NO. 1 over a region of at least 20, preferably at least 30, for instance 40, 60 or 100 more contiguous amino acids. Such a polypeptide homologue will be referred to a polypeptide of the present invention.

Generally, a fragment of polypeptide comprising amino acid sequence shown in SEQ ID NO. 1, 4, 6, 9, 11 or 14 or its homologues will be at least 10, preferably at least 15, for example 20, 25, 30, 40, 50 or 60 amino acids in length, and are also referred to by the term "a polypeptide of the present invention".

A cDNA capable of selectively hybridizing to the cDNA comprising nucleotide sequence shown in SEQ ID NO. 2, 5, 7, 10, 12 or 15 will be generally at least 70%, preferably at least 80 or 90% and more preferably at least 95% homologous to the cDNA of SEQ ID NO. 2, 5, 7, 10, 12 or 15 over a region of at least 20, preferably at least 30, for instance 40, 60 or 100 or more contiguous nucleotides. Such cDNA will be referred to "a cDNA of the present invention".

Fragments of the cDNA comprising nucleotide sequence shown in SEQ ID NO. 2, 5, 7, 10, 12 or 15 will be at least 10, preferably at least 15, for example 20, 25, 30 or 40 nucleotides in length, and will be also referred to "a cDNA of the present invention" as used herein.

A further embodiment of the present invention provides replication and expression vectors carrying cDNA of the invention. The vectors may be, for example, plasmid, virus or phage vectors provided with an origin of replication, optionally a promoter for the expression of the said cDNA and optionally a regulator of the promoter. The vector may contain one or more selectable marker genes, for example a ampicillin resistance gene. The

vector may be used in vitro, for example of the production of RNA corresponding to the cDNA, or used to transfect or transfet a host cell.

A further embodiment of the present invention provides host cells transformed with the vectors for the replication and expression of the cDNA of the invention, including the nucleotide sequence shown in SEQ ID NO. 2, 3, 5, 7, 8, 10, 12, 13 or 15 or the open reading frame thereof. The cells will be chosen to be compatible with the vector and may for example be bacterial, yeast, insect or mammalian.

A further embodiment of the present invention provides a method of producing a polypeptide which comprises culturing host cells of the present invention under conditions effective to express a polypeptide of the invention. Preferably, in addition, such a method is carried out under conditions in which the polypeptide of the invention is expressed and then produced from the host cells.

cDNA of the present invention may also be inserted into the vectors described above in an antisense orientation in order to prove for the production of antisense RNA. Such antisense RNA may be used in a method of controlling the levels of a polypeptide of the invention in a cell.

The invention also provides monoclonal or polyclonal antibodies against a polypeptide of the invention. The invention further provides a process for the production of monoclonal or polyclonal antibodies to the polypeptides of the invention. Monoclonal antibodies may be prepared by common hybridoma technology using polypeptides of the invention or fragments thereof, as an immunogen. Polyclonal antibodies may also be prepared by common means which comprise inoculating host animals, for example a rat or a rabbit, with polypeptides of the invention and recovering immune serum.

The present invention also provides pharmaceutical compositions containing a polypeptide of the invention, or an antibody thereof, in

association with a pharmaceutically acceptable diluent and/or carrier.

The polypeptide of the present invention includes that which a part of their amino acid sequence is lacking (e.g., a polypeptide comprised of the only essential sequence for revealing a biological activity in an amino acid sequence shown in SEQ ID NO.1), that which a part of their amino acid sequence is replaced by other amino acids (e.g., those replaced by an amino acid having a similar property) and that which other amino acids are added or inserted into a part of their amino acid sequence, as well as those having the amino acid sequence shown in SEQ ID NO. 1, 4, 6, 9, 11 or 14.

As known well, there are one to six kinds of codon as that encoding one amino acid (for example, one kind of codon for Methioine (Met), and six kinds of codon for leucine (Leu) are known). Accordingly, the nucleotide sequence of cDNA can be changed in order to encode the polypeptide having the same amino acid sequence.

The DNA of the present invention, specified in (2) includes a group of every nucleotide sequences encoding polypeptides (1) shown in SEQ ID NO. 1, 4, 6, 9, 11 or 14. There is a probability that yield of a polypeptide is improved by changing a nucleotide sequence.

The cDNA specified in (3) is the embodiment of the cDNA shown in (2), and indicate the sequence of natural form.

The cDNA shown in (4) indicates the sequence of the cDNA specified in (3) with natural non-translational region.

cDNA carrying nucleotide sequence shown in SEQ ID NO. 3, 8 or 13 is prepared by the following method:

Brief description of Yeast SST method (see USP No. 5,536,637) is as follows.

Yeast such as *Saccharomyces cerevisiae* should secrete invertase into the medium in order to take sucrose or raffinose as a source of energy or

carbon.

(Invertase is an enzyme to cleave raffinose into sucrose and melibiose, sucrose into fructose and glucose.)

It is known that many known mammalian signal peptide make yeast secrete its invertase.

From these knowledge, SST method was developed as a screening method to find novel signal peptide which make it possible can to secrete yeast invertase from mammalian cDNA library. SST method uses yeast growth on raffinose medium as a marker. Non-secretory type invertase gene SUC2 (GENBANK Accession No. V 01311) lacking initiation codon ATG was inserted to yeast expression vector to prepare yeast SST vector pSUC2.

In this expression vector, ADH promoter, ADH terminator (both were derived from AAH5 plasmid (Gammerer, Methods in Enzymol. 101, 192-201, 1983)), 2u ori (as a yeast replication origin), TRP1 (as a yeast selective marker), ColE1 ori(as a E. Coli replication origin) and ampicillin resistance gene (as a drug resistance marker) were inserted.

Mammalian cDNA was inserted into the upstream of SUC2 gene to prepare yeast SST cDNA library. Yeast lacking secretory type invertase, was transformed with this library.

If inserted mammalian cDNA encodes a signal peptide, yeast could be survive in raffinose medium as a result of restoring secretion of invertase. Only to culture yeast colonies, prepare plasmids and determine the nucleotide sequence of the insert cDNAs, it is possible to identify novel signal peptide rapidly and easily.

Preparation of yeast SST cDNA library is as follows:

(1) mRNA is isolated from the targeted cells, second-strand synthesis is performed by using random primer with certain restriction enzyme (enzyme I) recognition site,

- (2) double-strand cDNA is ligated to adapter containing certain restriction endonuclease (enzyme II) recognition site, differ from enzyme I, digested with enzyme I and fractionated in a appropriate size,
- (3) obtained cDNA fragment is inserted into yeast expression vector on the upstream region of invertase gene which signal peptide is deleted and the library was transformed.

Detailed description of each step is as follows:

- (1) mRNA is isolated from mammalian organs and cell lines stimulate them with appropriate stimulator if necessary by known methods (Molecular Cloning (Sambrook, J., Fritsch, E. F. and Maniatis, T., Cold Spring Harbor Laboratory Press, 1989) or Current Protocol in Molecular Biology (F. M. Ausubel et al, John Wiley & Sons, Inc.) if not remark especially).

Mouse embryonic heart is chosen as a tissue source. Double-strand cDNA synthesis using random primer is performed by known methods.

Any sites may be used as restriction endonuclease recognition site I which is linked to adapter and restriction endonuclease recognition site II which is used in step (2), if both sites are different each other. Preferably, XhoI is used as enzyme I and EcoRI as enzyme II.

In step (2), cDNA is created blunt-ends with T4 DNA polymerase, ligated enzyme II adapter and digested with enzyme I. Fragment cDNA is analyzed with agarose-gel electrophoresis and is selected cDNA fraction ranging in size from 300 to 800 bp. As mentioned above, any enzyme may be used as enzyme II if it is not same the enzyme I.

In step (3), cDNA fragment obtained in step (2) is inserted into yeast expression vector on the upstream region of invertase gene which signal peptide is deleted. E. coli transformed with the expression vector. Many vectors are known as yeast expression plasmid vector. For example, YEp24 is also functioned in E. Coli. Preferably pSUC2 as described above is used.

Many host E. Coli strains are known for transformation, preferably DH10B competent cell is used. Any known transformation method is available, preferably it is performed by electroporation method. Transformant is cultured by known methods to obtain cDNA library for yeast SST method.

However not every All of the clones do not contain cDNA fragment. Further all of the gene fragments do not encode unknown signal peptides. It is therefore necessary to screen a gene fragment encoding for an unknown signal peptide from the library.

Therefore, screening of fragments containing a sequence encoding an appropriate signal peptide is performed by transformation of the cDNA library into *Saccharomyces cerevisiae* (e.g. YT455 strain) which lack invertase (it may be prepared by known methods.).

Transformation of yeast is performed by known methods, e.g. lithium acetate method. Transformant is cultured in a selective medium, then transferred to a medium containing raffinose as a carbon source. Survival colonies are selected and then prepared plasmid. Survival colonies on a raffinose-medium indicates that some signal peptide of secretory protein was inserted to this clone.

Isolated positive clones is determined the nucleotide sequence. As to a cDNA encodes unknown protein, full-length clone may be isolated by using cDNA fragment as a probe and then determined to obtain full-length nucleotide sequence. These manipulation is performed by known methods.

Once the nucleotide sequences shown in SEQ ID NO. 2, 5, 7, 10, 12 or 15 are determined partially or preferably fully, it is possible to obtain cDNA encode mammalian protein itself, homologue or subset of the invention. cDNA library or mRNA derived from mammals was screened by PCR with any synthesized oligonucleotide primers or by hybridization with any fragment as a probe. It is possible to obtain cDNA encodes other mammalian homologue

protein from other mammalian cDNA or genome library.

If a cDNA obtained above contains a nucleotide sequence of cDNA fragment obtained by SST (or concensus sequence thereof), it will be thought that the cDNA encodes signal peptide. So it is clear that the cDNA will be full-length or almost full.

(All signal peptides exist at N-termini of a protein and are encoded at 5'-temini of open reading frame of cDNA.)

The confirmation may be carried out by Northern analysis with the said cDNA as a probe. It is thought that the cDNA is almost complete length, if length of the cDNA is almost the same length of the mRNA obtained in the hybridizing band.

The present invention supplies full-length protein and also its mature protein sequence. The full-length protein sequence deduced from nucleotide sequences shown in SEQ ID NO. 2, 7 or 12.

Mature proteins are obtained by expressing full-length cDNAs shown in SEQ ID NO. 3, 8 or 13 in mammalian cells or other host cells.

Mature protein sequences are deduced from their full-length amino acid sequences.

Once the nucleotide sequences shown in SEQ ID NOS. 2, 5, 7, 10, 12 or 15 are determined, cDNAs of the present invention are obtained by chemical synthesis, or by hybridization making use of nucleotide fragments which are chemically synthesized as a probe.

Furthermore, cDNAs of the present invention are obtained in desired amount by transforming a vector that contains the cDNA into a proper host, and culturing the transformant.

The polypeptides of the present invention may be prepared by:

- (1) isolating and purifying from an organism or a cultured cell,
- (2) chemically synthesizing, or

(3) using recombinant DNA technology,  
preferably, by the method described in (3) in industrial production.

Examples of expression system for (host-vector system) producing a polypeptide by using recombinant DNA technology are the expression systems of bacteria, yeast, insect cells and mammalian cells.

In the expression of the polypeptide, for example, in E. Coli, the expression vector is prepared by adding the initiation codon (ATG) to 5' end of a DNA encoding mature peptide, connecting the DNA thus obtained to the downstream of a proper promoter (e.g., trp promoter, lac promoter, λ PL promoter, T7 promoter etc.), and then inserting it into a vector (e.g., pBR322, pUC18, pUC19 etc.) which functions in an E. coli strain.

Then, an E. coli strain (e.g., E. coli DH1 strain, E. coli JM109 strain, E. coli HB101 strain, etc.) which is transformed with the expression vector described above may be cultured in an appropriate medium to obtain the desired polypeptide. When a signal peptide of bacteria (e.g., signal peptide of pel B) is utilized, the desired polypeptide may be also released in periplasm. Furthermore, a fusion protein with other polypeptide may be also produced easily.

In the expression of the polypeptide, for example, in a mammalian cells, for example, the expression vector is prepared by inserting the DNA encoding cDNA shown in SEQ ID NO. 3, 8 or 13 into the downstream of a proper promoter (e.g., SV40 promoter, LTR promoter, metallothionein promoter etc.) in a proper vector (e.g., retrovirus vector, papilloma virus vector, vaccinia virus vector, SV40 vector, etc.) a proper mammalian cell (e.g., monkey COS-7 cell, Chinese hamster CHO cell, mouse L cell etc.) is transformed with the expression vector thus obtained, and then the transformant is cultured in a proper medium to get a desired polypeptide in the culture medium. Further, fusion protein may be produced by linking cDNA fragment encoding other

polypeptide such as Fc portion of an antibody. The polypeptide thus obtained may be isolated and purified by conventional biochemical methods.

#### **Effect of the invention**

The polypeptides of the present invention and cDNA encoding them are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified below.

Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or by administration or use of cDNA encoding them (such as, for example, in gene therapies or vectors suitable for introduction of DNA).

We have been confirmed that the said polypeptide possess the suppressing activity on the differentiation of vascular smooth muscle cells.

Accordingly, the polypeptides may be useful for treatment of diseases related to abnormal proliferation of smooth muscle cells, for example, arteriosclerotic coronary disease, neointimal formation which results in restenosis after percutaneous transluminal coronary angioplasty and myosarcoma.

But not limit the present invention:

<Cytokine activity and cell proliferation/differentiation activity>

The protein of the present invention may exhibit cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations.

Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity.

The activity of a protein of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines.

<Immune stimulating/suppressing activity>

The protein of the present invention may also exhibit immune stimulating or immune suppressing activity. The protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations.

These immune deficiencies may be genetic or be caused by viral (e.g. HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders.

More specifically, infectious diseases causes by viral, bacterial, fungal or other infection may be treatable using the protein of the present invention, including infections by HIV, hepatitis viruses, herpes viruses, mycobacteria, leshmania, malaria and various fungal infections such as candida.

Of course, in this regard, a protein of the present invention may also be useful where a boost to the immune system generally would be indicated, i.e., in the treatment of cancer.

Such a protein of the present invention may also be useful in the treatment of allergic reactions and conditions, such as asthma or other respiratory problems.

The protein of the present invention may also suppress chronic or acute inflammation, such as, for example, that associated with infection (such as septic shock or systemic inflammatory response syndrome (SIRS)), inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1 (such as the effect demonstrated by IL-

11).

<Hematopoiesis regulating activity>

The protein of the present invention may be useful in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell deficiencies.

Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis.

The said biological activities are concerned with the following all or some example(s).

e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility. for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells;

in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression;

in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions;

and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal

hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either in-vivo or ex-vivo (i.e. in conjunction with bone marrow transplantation) as normal cells or genetically manipulated for gene therapy.

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

The activity of a protein of the invention may, among other means. be measured by the following methods :

<Tissue generation/regeneration activity>

The protein of the present invention also may have utility in compositions used for bone, cartilage, tendon, Ligament and/or nerve tissue growth or regeneration, as well as for wound healing and tissue repair, and in the treatment of burns, incisions and ulcers.

The protein of the present invention, which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals.

Such a preparation employing the protein of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints.

De novo bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

The protein of this invention may also be used in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce differentiation of progenitors of bone-forming cells.

The protein of the invention may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

Another category of tissue regeneration activity that may be attributable to the protein of the present invention is tendon/ligament formation. A protein of the present invention, which induces tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals.

Such a preparation employing a tendon/Ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue.

De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide an environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon Ligament cells or progenitors ex vivo for return in vivo to effect tissue repair.

The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering

agent as a carrier as is well known in the art.

The protein of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue. i.e. for the treatment of central and peripheral nervous system diseases and neuropathies. as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a protein may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome.

Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

It is expected that the protein of the present invention may also exhibit activity for generation of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting or suppressing the proliferation of cells comprising such tissues. Part of the desired effects may be by inhibition of fibrotic scarring to allow normal tissue to regenerate.

A protein of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.  
**<Activin/Inhibin activity>**

The protein of the present invention may also exhibit activin- or

inhibin-related activities. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH).

Thus, a protein of the present invention, alone or in heterodimers with a member of the inhibin \*a family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals.

Alternatively, the protein of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin-\*b group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary.

See for example, USP 4,798,885. The polypeptide of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as cows, sheep and pigs.

<Chemotactic/chemokinetic activity>

A protein of the present invention may have chemotactic or chemokinetic activity (e.g., act as a chemokine) for mammalian cells, including, for example, monocytes, neutrophils, T-cells, mast cells, eosinophils and/or endothelial cells.

Chemotactic and chemokinetic proteins can be used to mobilized or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic proteins provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the

tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

<Hemostatic and thrombolytic activity>

The protein of the invention may also exhibit hemostatic or thrombolytic activity. As a result, such a protein is expected to be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A protein of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction or stroke).

<Receptor/ligand activity>

The protein of the present invention may also demonstrate activity as receptors, receptor ligands or inhibitors or agonists of receptor/ligand interactions. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses). Receptors and ligands are also useful for screening of potential peptide

or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

<Nutritional uses>

Proteins of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the protein of the present invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the protein of the invention can be added to the medium in or on which the microorganism is cultured.

<Cadherin/Tumor invasion suppresser activity>

Cadherins are calcium-dependent adhesion molecules that appear to play major roles during development, particularly in defining specific cell types. Loss or alteration of normal cadherin expression can lead to changes in cell adhesion properties linked to tumor growth and metastasis.

Cadherin malfunction is also implicated in other human diseases, such as pemphigus vulgaris and pemphigus foliaceus (autoimmune blistering skin diseases), Crohn's disease, and some developmental abnormalities.

The cadherin superfamily includes well over forty members, each with a distinct pattern of expression. All members of the superfamily have in common conserved extracellular repeats (cadherin domains), but structural differences are found in other parts of the molecule.

The cadherin domains bind calcium to form their tertiary structure and thus calcium is required to mediate their adhesion. Only a few amino acids in

the first cadherin domain provide the basis for homophilic adhesion; modification of this recognition site can change the specificity of a cadherin so that instead of recognizing only itself, the mutant molecule can now also bind to a different cadherin. In addition, some cadherins engage in heterophilic adhesion with other cadherin.

E-cadherin, one member of the cadherin superfamily, is expressed in epithelial cell types. Pathologically, if E-cadherin expression is lost in a tumor, the malignant cells become invasive and the cancer metastasizes. Transfection of cancer cell line with cDNAs expressing E-cadherin has reversed cancer-associated changes by returning altered cell shapes to normal, restoring cells adhesiveness to each other and to their substrate, decreasing the cell growth rate, and drastically reducing anchorage-independent cell growth.

Thus, reintroducing E-cadherin expression reverts carcinomas to a less advanced stage. It is likely that other cadherins have the same invasion suppresser role in carcinomas derived from other tissue types. Therefore, proteins of the present invention with cadherin activity, and cDNAs of the present invention encoding such proteins, can be used to treat cancer. Introducing such proteins or cDNAs into cancer cells can reduce or eliminate the cancerous change observed in these cells by providing normal cadherin expression.

Cancer cells have also been shown to express cadherins of a different tissue type than their origin, thus allowing these cells to invade and metastasize in a different tissue in the body. Proteins of the present invention with cadherin activity, and cDNAs of the present invention encoding such proteins, can be substituted in these cells for the inappropriately expressed cadherins, restoring normal cell adhesive properties and reducing or eliminating the tendency of the cells to metastasize.

Additionally, proteins of the present invention with cadherin activity, and cDNA of the present invention encoding such proteins, can be used to generate antibodies recognizing and binding to cadherins.

Such antibodies can be used to block the adhesion of inappropriately expressed tumor-cell cadherins, preventing the cells from forming a tumor elsewhere. Such an anti-cadherin antibody can also be used as a marker for the grade, pathological type, and prognosis of a cancer, i.e. the more progressed the cancer, the less cadherin expression there will be, and this decrease in cadherin expression can be detected by the use of a cadherin-binding antibody.

Fragments of proteins of the present invention with cadherin activity, preferably a polypeptide comprising a decapeptide of the cadherin recognition site, and cDNAs of the present invention encoding such protein fragments, can also be used to block cadherin function by binding to cadherins and preventing them from binding in ways that produce undesirable effects. Additionally, fragments of proteins of the present invention with cadherin activity, preferably truncated soluble cadherin fragments which have been found to be stable in the circulation of cancer patients, and polynucleotides encoding such protein fragments, can be used to disturb proper cell-cell adhesion.

#### <Tumor Inhibiting activity>

In addition to the activities described above for immunological treatment or prevention of tumors, the protein of the invention may exhibit other anti-tumor activities. The protein may inhibit tumor growth directly or indirectly (such as, for example, via ADCC).

The protein may exhibit its tumor inhibitory activity by acting on tumor tissue or tumor precursor tissue, by inhibiting formation of tissues necessary to support tumor growth (such as, for example, by inhibiting angiogenesis), by causing production of other factors, agents or cell types

which inhibit tumor growth, or by suppressing, eliminating or inhibiting factors, agents or cell types which promote tumor growth.

<Other activity>

The protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, bacteria, viruses, fungi and other parasites;

effecting (suppressing or enhancing) bodily characteristics, including, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution);

effecting elimination of dietary fat, protein, carbohydrate;

effecting behavioral characteristics, including appetite, libido, stress, cognition (including cognitive disorders), depression and violent behaviors;

providing analgesic effects or other pain reducing effects;

promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages;

in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases.

The polypeptide with above activities, is suspected to have following functions by itself or interaction with its ligands or receptors or association with other molecules. For example, proliferation or cell death of B cells, T cells and/or mast cells or class specific induction of B cells by promotion of class switch of immunoglobulin genes; differentiation of B cells to antibody-forming cells; proliferation, differentiation, or cell death of precursors of granulocytes; proliferation, differentiation, or cell death of precursors of monocytes-macrophages;

proliferation, or up regulation or cell death of neutrophils, monocytes-macrophages, eosinophils and/or basophils; proliferation, or cell death of precursors of megakaryocytes; proliferation, differentiation, or cell death of precursors of neutrophils; proliferation, differentiation, or cell death of precursors of T cells and B cells; promotion of production of erythrocytes; sustainment of proliferation of erythrocytes, neutrophils, eosinophils, basophils, monocytes-macrophages, mast cells, precursors of megakaryocyte ; promotion of migration of neutrophils, monocytes-macrophages, B cells and/or T cells; proliferation or cell death of thymocytes; suppression of differentiation of adipocytes; proliferation or cell death of natural killer cells; proliferation or cell death of hematopoietic stem cells; suppression of proliferation of stem cells and each hematopoietic precursor cells; promotion of differentiation from mesenchymal stem cells to osteoblasts or chondrocytes, proliferation or cell death of mesenchymal stem cells, osteoblasts or chondrocytes and promotion of bone absorption by activation of osteoclasts and promotion of differentiation from monocytes to osteoclasts.

This peptide is also suspected to function to nervous system, so expected to have functions below; differentiation to kinds of neurotransmitter-responsive neurons, survival or cell death of these cells; promotion of proliferation or cell death of glial cells; spread of neural dendrites; survival or cell death of gangriocytes; proliferation, promotion of differentiation, or cell death of astrocytes; proliferation or survival of peripheral neurons; proliferation or cell death of Schwann cells; proliferation, survival or cell death of motoneurons.

Furthermore, in the process of development of early embryonic, this polypeptide is expected to promote or inhibit the organogenesis of epidermis, brain, backbone, and nervous system by induction of ectoderm, that of

notochord connective tissues(bone, muscle, tendon), hemocytes, heart, kidney, and genital organs by induction of mesoderm, and that of digestive apparatus (stomach, intestine, liver, pancreas), respiratory apparatus (lung, trachea) by induction of endoderm. In adult, also, this polypeptide is thought to proliferate or inhibit the above organs.

Therefore, this polypeptide itself is expected to be used as an agent for the prevention or treatment of disease of progression or suppression of immune, nervous, or bone metabolic function, hypoplasia or overgrowth of hematopoietic cells: inflammatory disease (rheumatism, ulcerative colitis, etc.), decrease of hematopoietic stem cells after bone marrow transplantation, decrease of leukocytes, platelets, B-cells, or T-cells after radiation exposure or chemotherapeutic dosage against cancer or leukemia, anemia, infectious disease, cancer, leukemia, AIDS, bone metabolic disease(osteoporosis etc.), arteriosclerosis, various degenerative disease (Alzheimer's disease, multiple sclerosis, etc.), or nervous lesion.

In addition, since this polypeptide is thought to induce the differentiation or growth of organs derived from ectoderm, mesoderm, and endoderm, this polypeptide is expected to be an agent for tissue repair (epidermis, bone, muscle, tendon, heart, kidney, stomach, intestine, liver, pancreas, lung, and trachea, etc.).

Quantitation of this polypeptide in the body can be performed using polyclonal or monoclonal antibodies against this polypeptide. It can be used the study of relationship between this polypeptide and disease or diagnosis of disease, and so on. Polyclonal and monoclonal antibodies can be prepared using this polypeptide or its fragment as an antigen by known method.

Identification, purification or molecular cloning of known or unknown proteins which bind this polypeptide can be performed using this polypeptide by, for example, preparation of the affinity-column.

Identification of the molecules which interact with this polypeptide and molecular cloning of the gene can be performed by west-western method using this polypeptide or by yeast two-hybrid system using the cDNA (preferably cDNA encoding this polypeptide).

Agonists/antagonists of this receptor polypeptide and inhibitors between receptor and signal transduction molecules can be screened using this polypeptide.

For example, the screening can be carried out the following method.

a) The reaction mixtures, which contain this polypeptide, screening compound and the cells, are incubated under the condition which the cells are normally stimulated by this peptide. (The reaction mixtures also contain the labeled compound, which is introduced into the cells according to the cell proliferation, and peptide which allow to observe the function of this peptide efficiently.)

b) Decision that the compounds are efficient agonists/antagonists or not, are performed by measurement of cell proliferation ability.

More detailed methods are followed:

Rat vascular muscle cell line (ATCC CRL-1444 or CRL1476) is cultured in 96 well plate with 10%FBS for 24 hours. Then the culture medium are replaced to the serum-free medium supplemented with each several concentrations of human PDGF-BB. At that time compounds to screen as well as A55 protein are added in the medium when screening the antagonists of A55 protein. While, compounds alone are added in the medium when screening the agonists of A55 protein. After 24 hours incubation, these cells are pulsed for 4hours with  $^3\text{H}$ -thymidine. By measuring the  $^3\text{H}$ -thymidine incorporation, it is possible to determine whether the compounds have inhibitory or stimulatory effect on the A55 activity.

cDNAs of the present invention are useful not only the important and

essential template for the production of the polypeptide of the present invention which is expected to be largely useful, but also be useful for diagnosis or therapy (for example, treatment of gene lacking, treatment to stop the expression of the polypeptide by antisense DNA (RNA)). Genomic DNA may be isolated with the cDNA of the present invention, as a probe. As the same manner, a human gene encoding which can be highly homologous to the cDNA of the present invention, that is, which encodes a polypeptide highly homologous to the polypeptide of the present invention and a gene of animals excluding mouse which can be highly homologous to the cDNA of the present invention, also may be isolated.

#### **Application for Pharmaceuticals**

For the medical treatment for diseases described above, the polypeptide of the invention or the antibody of the polypeptide of the invention may be administered systemically or partially in most cases, usually by oral or parenteral administration, preferably orally, intravenously or intraventricularly.

The doses to be administered depend upon age, body weight, symptom, desired therapeutic effect, route of administration, and duration of the treatment etc. In human adults, one dose per person is generally between 100 µg and 100 mg, by oral administration, up to several times per day, and between 10 µg and 100 mg, by parenteral administration up to several times per day.

As mentioned above, the doses to be used depend upon various conditions. Therefore, there are cases in which doses lower than or greater than the ranges specified above may be used.

The compounds of the present invention, may be administered as solid compositions, liquid compositions or other compositions for oral

administration, as injections, liniments or suppositories etc. for parenteral administration.

Solid compositions for oral administration include compressed tablets, pills, capsules, dispersible powders, granules. Capsules include soft or hard capsules.

In such compositions, one or more of the active compound(s) is or are admixed with at least one inert diluent (such as lactose, mannitol, glucose, hydroxypropyl cellulose, microcrystalline cellulose, starch, polyvinylpyrrolidone, magnesium metasilicate aluminate, etc.). The compositions may also comprise, as is normal practice, additional substances other than inert diluents: e.g. lubricating agents (such as magnesium stearate etc.), disintegrating agents (such as cellulose calcium glycolate, etc.), stabilizing agents (such as human serum albumin, lactose etc.), and assisting agents for dissolving (such as arginine, asparagine acid etc.).

The tablets or pills may, if desired, be coated with a film of gastric or enteric materials (such as sugar, gelatin, hydroxypropyl cellulose or hydroxypropylmethyl cellulose phthalate, etc.), or be coated with more than two films. And then, coating may include containment within capsules of absorbable materials such as gelatin.

Liquid compositions for oral administration include pharmaceutically-acceptable emulsions, solutions, syrups and elixirs. In such compositions, one or more of the active compound(s) is or are contained in inert diluent(s) commonly used (purified water, ethanol etc.). Besides inert diluents, such compositions may also comprise adjuvants (such as wetting agents, suspending agents, etc.), sweetening agents, flavoring agents, perfuming agents, and preserving agents.

Other compositions for oral administration include spray compositions which may be prepared by known methods and which comprise one or more of

the active compound(s). Spray compositions may comprise additional substances other than inert diluents: e.g. stabilizing agents (sodium sulfite etc.), isotonic buffer (sodium chloride, sodium citrate, citric acid, etc.). For preparation of such spray compositions, for example, the method described in the United States Patent No. 2,868,691 or 3,095,355 (herein incorporated in their entireties by reference) may be used.

Injections for parenteral administration include sterile aqueous or non-aqueous solutions, suspensions and emulsions. In such compositions, one or more active compound(s) is or are admixed with at least one inert aqueous diluent(s) (distilled water for injection, physiological salt solution, etc.) or inert non-aqueous diluents(s)(propylene glycol, polyethylene glycol, olive oil, ethanol, POLYSOLBATE 80 TM , etc.).

Injections may comprise additional compound other than inert diluents: e.g. preserving agents, wetting agents, emulsifying agents, dispersing agents, stabilizing agent (such as human serum albumin, lactose, etc.), and assisting agents such as assisting agents for dissolving (arginine, asparaginic acid, etc.).

**Examples:**

The following examples concerning clone A55 are illustrated, but not limit the present invention.

**Example 1**

**Preparation of poly(A)+RNA**

Total RNA was prepared from mouse day18.5 embryonic heart by TRIzol reagent (Trade Mark, GIBCOBRL), and poly (A)<sup>+</sup> RNA was purified from the total RNA by mRNA Purification Kit (Trade Mark, Pharmacia).

## Example 2

### Preparation of yeast SST cDNA library

Double strand cDNA was synthesized by SuperScript Plasmid System for cDNA Synthesis and Plasmid Cloning (brand name, GIBCOBRL) with above poly(A)+RNA as template and random 9mer as primer which was containing XhoI site:

5'-CGA TTG AAT TCT AGA CCT GCC TCG AGN NNN NNN NN-3'

cDNA was ligated EcoRI adapter by DNA ligation kit ver.2 (trade name, Takara Shuzo; this kit was used in all ligating steps hereafter.) and digested by XhoI. cDNAs were separated by agarose-gel electrophoresis. 300 - 800 bp cDNAs were isolated and were ligated to EcoRI/NotI site of pSUC2 (see US 5,536,637). E. Coli DH10B strain were transformed by pSUC2 with electroporation to obtain yeast SST cDNA library.

## Example 3

### Screening by SST method and DNA sequencing of positive clone

Plasmids of the cDNA library were prepared. Yeast YTK12 strain were transformed by the plasmids with lithium acetate method (Current Protocols In Molecular Biology 13.7.1). The transformed yeast were plated on triptphan-free medium (CMD-Try medium) for selection. The plate was incubated for 48 hour at 30 oC. Replica of the colony which is obtained by Accutran Replica Plater (trade name, Schleicher & Schuell) were place YPR plate containing raffinose for carbon source, and the plate was incubated for 14 days at 30 oC.

After 3 days, each colony appeared was streaked on YPR plate again. The plates were incubated for 48 hours at 30 oC. Single colony was inoculated to YPR medium and was incubated for 48 hours at 30 oC. Then plasmids were prepared.

Insert cDNA was amplified by PCR with two kind primers which exist end side of cloning site on pSUC2 (sense strand primers were biotinylated). Biotinylated single strand of cDNAs were purified with Dynabeads (trade name, DYNAL) and determined the nucleotide sequences.

Sequencing was performed by Dye Terminator Cycle Sequencing Ready Reaction with DNA Sequencing kit (trade name, Applied Biosystems Inc.) and sequence was determined by DNA sequencer 373 (Applied Biosystems Inc.). All sequencing hereafter was carried with this method.

The clone named A55 is not registered on databases by homology search of cDNA sequence and deduced amino acid sequence and so it is cleared that the sequence is novel one. So, we tried to isolate clone full-length cDNA of the fragment of A55 clone (hereafter A55 SST fragment cDNA). We confirmed that A55 SST fragment cDNA contains signal peptide by comparison with known peptide which has signal peptide in view of function and structure.

#### Example 4

##### Cloning and sequencing of a full-length cDNA of A55

Phage particles of a cDNA library of mouse day13 embryonic heart(uni-ZAP XR, Stratagene) were transfected to E. coli XL1-Blue MRF\* host cells (Stratagene). Obtained one million plaques were transferred to nylon membranes. The membranes were hybridized with <sup>32</sup>P-labeled mouse A55 SST fragment cDNA as a probe. Many positive plaques were obtained.

From one positive plaque the phage particles containing a cloned insert were prepared, and were subjected to conversion into phagemid particles (pBluescript SK(-)) by co-infection of E. coli XL1-Blue MRF\* host cells(Stratagene) with ExAssist helper phage(Stratagene). The phagemid particles were transfected to E. coli DH5a. The plasmids were prepared from the obtained transformants.

Nucleotide sequence of 5'-end of cDNA were determined to confirm the existence

of the sequences of SST fragment cDNA. And then full-length sequencing were performed to obtain SEQ ID NO.3.

An open reading frame was determined and translation region for amino acid sequence shown in SEQ ID NO. 2 and deduced full-length amino acid sequence shown in SEQ ID NO. 1 were obtained. Mature protein of the said polypeptide was deduced to 425 amino acids shown in SEQ ID NO. 3 (144..1418) or 423 amino acids shown in SEQ ID NO. 4. Translation region of SEQ ID NO. 4 is shown in SEQ ID NO. 5.

It was confirmed that there was no identical sequences to the DNA of the present invention by homology search program, BLASTN and FASTA against public nucleotide database. And it was also confirmed that there were no identical sequences to the polypeptide of the present invention (mouse A55 protein) by homologue search program, BLASTX, BLASTP and FASTA against amino acid database.

It is revealed that the polypeptide of the present invention, mouse A55 is novel secretion protein since the polypeptide have no trans-membrane region by hydrophobisity analysis of the amino acid sequence.

It was revealed that A55 protein contained six EGF like domains by motif search, so it was expected that clone A55 also possesses EGF family like activities. Significant homology were also recognized between the amino acid sequence of clone mouse A55 (1-448 AA region) and the one of human S1-5 (SwissProt Accession No. HSU03877) (1-387 AA region) by the comparison using BLASTX, BLASTP and FASTA. It was reported that human S1-5 was a secreted protein containing EGF like domain, was induced in fibroblasts by growth arrest, and stimulated DNA synthesis (Beata Lecka-Czernik et. al. Mol. Cell. Biol. 15, 120-128, 1995). Farther it was revealed that A55 protein was homologous to many proteins containing EGF-like domain.

**Example 5**

**Isolation of isoform gene of mouse A55 protein**

Initiation codon was determined by cloning of 5'-end cDNA by 5'-RACE (Rapid Amplification of cDNA Ends method using Marathon cDNA Amplification Kit (trade name, Clontech). Double stranded cDNA template was prepared from poly(A)+RNA of mouse embryonic heart tissue. Primer mA55-R1:

5' - CGT TTG TGC ACT GCT GCT GTG CAT TCC -3'

was prepared based on the information of full-length nucleotide sequences. PCR was performed with the said primer and adapter primer attached in the kit.

Amplified cDNA was separated with agarose-gel electrophoresis, and to pGEM-T Vector (trade name, Promega), ligated in and transformed to E. Coli DH5a and then plasmid was prepared. The full-length nucleotide sequences were determined. We found two different 5'-end sequences. One was identical to the clone containing the sequence in SEQ ID NO. 3, the other contained unknown sequence and no translational start site ATG (See SEQ ID NO. 7 and 8). The region defined from exon 1 of the clone was replaced by another exon which exists 400 bp downstream region of exon 1 was clarified by gene analysis. So it was cleared that the clone shown in SEQ ID NO. 8 was generated by alternative splicing of exon 1.

The clone encodes isoform protein shown in SEQ ID NO. 6 (6 amino acids in N termini of SEQ ID NO. 1 was replaced by 19 amino acids in N termini of SEQ ID NO. 6).

The mature protein of this polypeptide was deduced 425 amino acids shown in SEQ ID NO. 8 (340...1614) or 423 amino acids shown in SEQ ID NO. 9. SEQ ID NO. 10 is the translational region of the polypeptide shown in SEQ ID

NO. 9.

Example 6

Determination of nucleotide sequence of human A55 gene

The present inventors found that Human EST sequence (GENBANK Accession No. H17726) homologous to 5'-end sequence of mouse A55 in the process of homology search shown in example 4.

And the present inventors buy the Clone ID 50483 derived from human brain cDNA library GENBANK Accession No. H17726 from American Type Culture Collection (ATCC). The full-length nucleotide sequence shown in SEQ ID NO. 13 was determined with the same manner as in the determination of mouse A55. Open reading frame was determined and translational region shown in SEQ ID NO. 12 and deduced amino acid sequence shown in SEQ ID NO. 11 were obtained. From above results, it is clarified that the human clone is full-length and have 89.3 % homology to mouse A55 at DNA level (translational region) and have 94.2 % homology to the one at amino acid level. It is suggested that the obtained human clone should be human counterpart of mouse A55.

(The clone was called human A55 hereafter.)

The mature protein of this polypeptide was deduced 425 amino acids shown in SEQ ID NO. 13 (238...1512) or 423 amino acids shown in SEQ ID NO. 14. Translational region of the polypeptide shown in SEQ ID NO. 14 shows in SEQ ID NO. 15.

It was confirmed that there was no identical sequences to the DNAs of the present invention by homology search program, BLASTN and FASTA against public nucleotide database. And it was also confirmed that there were no identical sequences to human A55 proteins by homologue search program, BLASTX, BLASTP and FASTA against amino acid database.

Example 7

Mouse A55 protein expression in mammalian cell

Mouse full-length cDNA shown in SEQ ID NO. 3 was inserted into expression vector for mammalian cell pNotS (Kaufman et al., Nucleic Acids Res. 19, 4485-4490 (1991)) and mouse A55 expression plasmid pNotS-mA55 was constructed.

293T cells (which is derived from 293 cells (ATCC CRL-1573) and it stably transfected with SV40 T antigen) were transfected with pNotS and pNotS-mA55 using lipofection (GIBCOBRL). After preincubated for 19 hours, the cells were pulsed for 30 minutes with  $^{35}\text{S}$ -Met in the Met-free medium. Then the cells were incubated in the medium containing Met for 5 hours. Supernatant of the cells was recovered and concentrated 10-fold using centricron-10 (trade name, AMICON). Samples were subjected to SDS-polyacrylamide-gel electrophoresis. The gel was dried and  $^{35}\text{S}$ -labeled proteins were detected with BAS 2000 (Fuji Film).

A band was detected at 60-70 kDa in the supernatant of pNotS-mA55-transfected 293T cells. This band was not detected in the supernatant of pNotS-transfected 293T cells. This results confirmed that recombinant mouse A55 protein was expressed and secreted into the medium.

Molecular weight (60-70 kDa) of this recombinant mouse A55 protein was greater than it (48 kDa) predicted from its amino acid sequences. As this protein had two potential N-linked glycosylation sites and many Ser and Thr residues in which O-linked glycosyl chain could be added , it was suggested that the mouse A55 protein was a glycoprotein.

Example 8

Measurement of inhibition on proliferation of rat vascular smooth muscle cells by mouse A55 protein

Vascular smooth muscle cells were isolated from rat aorta ranging from heart to diaphragm and cultured primarily by the methods described in Shin Seikagaku Jikken Kouza 10 (The Japanese Biochemical Society).

These cells were co-incubated with 1, 3 or 10 ng/ml of human recombinant PDGF-BB (Genzyme) and 10% (v/v) of the mock or mA55 supernatant prepared according to the method described in example 7. And BrdU incorporation was measured using a Cell Proliferation ELISA, BrdU colorimetric kit (Boehringer-Mannheim).

The supernatant from 293T cells transfected with pNotS-mA55 significantly inhibited BrdU incorporation of rat primary vascular smooth muscle cells, while the supernatant from 293T cells transfected with only pNotS show no effect as shown in Fig. 1.

Moreover the supernatant from 293T cells transfected with pNotS-mA55 also inhibited BrdU incorporation even when rat vascular smooth muscle cells were stimulated with 1, 3 or 10 ng/ml of PDGF and increased BrdU incorporation in a dose-dependent manner, whereas the supernatant from 293T cells transfected with only pNotS did not affect compared with no supernatant addition (See Fig. 1).

These data revealed that the recombinant mouse A55 protein had the growth inhibitory activity on vascular smooth muscle cells.

#### Example 9

Human A55 protein expression in mammalian cell

Human A55 expression plasmid, pNotS-hA55, was constructed by inserting human full-length cDNA shown in SEQ ID NO. 13 into was into expression vector for mammalian cell pNotS (Kaufman et al., Nucleic Acids Res. 19, 4485-4490 (1991)).

Cos1 cells were transfected with pNotS and pNotS-hA55 by lipofectin

(GIBCOBRL). After preincubation for 24 hours, the cells were pulsed for 5 hours with 35S-Met and 35S-Cys. Supernatant of the cells was recovered and concentrated to 10-fold using centricon-10 (trade name, AMICON). Proteins in concentrated supernatant were separated by electrophoresis through SDS-page. The gel was dried and 35S-labeled proteins were detected with BAS 2000 (Fuji Film).

A band was detected at 60-70 kDa in the supernatant from Cos1 cells transfected with pNotS-hA55. This band was not detected in the supernatant of pNotS transfected Cos1 cells. These results confirmed that recombinant human A55 protein was expressed and secreted into the medium. And human A55 protein was also suggested that sugar chains were also added to human A55 protein as well as mouse A55.

#### Example 10

Detection of the inhibitory activity on proliferation of rat vascular smooth muscle cells by human A55 protein

A DNA fragment encoding a signal sequence of honey bee mertin, a tag sequence of six His residues and an enterokinase cleavage site was added to 5'-end of human A55 cDNA sequence from 238 to 1515 in SEQ ID NO. 13 or sequence in SEQ ID NO. 15 followed by stop codon and inserted into expression vector pNotS. Cos1 cells were transfected with pNotS-hA55 plasmid DNA and pNotS control plasmid DNA. The supernatant was recovered , digested by enterokinase, pulled through nickel column to remove the linker peptide, and then concentrated 10-fold by centricon-10 (trade name, AMICON).

After rat vascular muscle cell line (ATCC CRL-1444) was cultured in 96 well plates with 10% FBS for 24 hours, these cells were incubated for 24 hours with serum-free medium supplemented with several concentrations (1, 10 or 10 ng/ml) of human PDGF-BB (Genzyme) and with 10 % total volume

of supernatant of Cos1 cells which were transfected with pNotS-hA55 or pNotS, and then were pulsed for 4 hours with 3H-thymidine. After harvesting 3H-thymidine incorporation was detected. In this cell line remarkable decrease of 3H-thymidine incorporation were observed by supplement with hA55 supernatant, while there was no effect in the presence of the control supernatant.

Moreover similar effects were also observed when using other rat vascular smooth muscle cell lines (ATCC CRL-1476 and CRL-2018) and human vascular smooth muscle cell line (ATCC CRL-1999). These results revealed that the recombinant human A55 protein also had the growth inhibitory activity on the vascular smooth muscle cells as well as mouse A55 protein.

Morphological change was observed on the vascular smooth muscle cells treated with the supernatant from hA55-transfected Cos1 cells by microscopy detection. While no morphological change was observed on melanoma cell line SK-MEL-28 at the same experiment. Furthermore, hA55 protein was observed to induce the expression of chemokine JE and JK.

#### Experiment 11

##### Preparation of anti mouse A55 polyclonal antibody

Three kinds of peptide fragments of mouse A55 were synthesized by solid phase method:

RTNPVYRGPYSNPYSTS YSG (71-90) (48-67 of SEQ ID NO. 1)

GAYYIFQIKSGNEGREFYMR (376-395) (353-372 of SEQ ID NO. 1)

MTRPIKGPRDIQLDLEMITVN (406-426) (383-403 of SEQ ID NO. 1).

Rabbits were immunized to these peptides as immunogen and the serum were prepared after measurement of the activity. Each anti-mouse A55

antibody was purified by affinity column immobilized each peptide which was used as immunogen from the obtained serum.

The supernatant prepared by the same method described in example 7, was subjected to SDS-PAGE, the separated proteins were transferred to Immobilon-P (PVDF membrane, trade name, Millipore) from the acrylamide gel.

After blocking the membranes they were incubated with the anti mouse A55 polyclonal antibody as the first antibody and by developing using ECL kit (Amersham), the recombinant mouse A55 protein was detected.

A 60 k Da band was detected in the supernatant from mA55 transfected Cos1 cells as well as 35S-labeling experiment described in example 7. While no bands were detected in the supernatant from mock-transfected Cos1 cells. These results confirmed that the obtained polyclonal antibodies specifically recognized the mouse A55 protein.

#### **Brief Description of Figures**

Fig. 1 It shows that mouse A55 protein inhibits proliferation of rat aortic vascular smooth muscle cells which was stimulated by PDGF.

Fig. 2 It shows that human A55 protein inhibits proliferation of rat aortic vascular smooth muscle cells which was stimulated by PDGF.

**Sequence Listing**

SEQ ID NO. : 1

Length : 448 amino acids

Type : amino acid

Topology : liner

Molecule type : protein

Sequence Description :

Met Pro Gly Leu Lys Arg Ile Leu Thr Val Thr Ile Leu Ala Leu Trp

-23 -20 -15 -10

Leu Pro His Pro Gly Asn Ala Gln Gln Gln Cys Thr Asn Gly Phe Asp

-5 1 5

Leu Asp Arg Gln Ser Gly Gln Cys Leu Asp Ile Asp Glu Cys Arg Thr

10 15 20 25

Ile Pro Glu Ala Cys Arg Gly Asp Met Met Cys Val Asn Gln Asn Gly

30 35 40

Gly Tyr Leu Cys Ile Pro Arg Thr Asn Pro Val Tyr Arg Gly Pro Tyr

45 50 55

Ser Asn Pro Tyr Ser Thr Ser Tyr Ser Gly Pro Tyr Pro Ala Ala Ala

60 65 70

Pro Pro Val Pro Ala Ser Asn Tyr Pro Thr Ile Ser Arg Pro Leu Val

75 80 85

Cys Arg Phe Gly Tyr Gln Met Asp Glu Gly Asn Gln Cys Val Asp Val  
90 95 100 105

Asp Glu Cys Ala Thr Asp Ser His Gln Cys Asn Pro Thr Gln Ile Cys  
110 115 120

Ile Asn Thr Glu Gly Gly Tyr Thr Cys Ser Cys Thr Asp Gly Tyr Trp  
125 130 135

Leu Leu Glu Gly Gln Cys Leu Asp Ile Asp Glu Cys Arg Tyr Gly Tyr  
140 145 150

Cys Gln Gln Leu Cys Ala Asn Val Pro Gly Ser Tyr Ser Cys Thr Cys  
155 160 165

Asn Pro Gly Phe Thr Leu Asn Asp Asp Gly Arg Ser Cys Gln Asp Val  
170 175 180 185

Asn Glu Cys Glu Thr Glu Asn Pro Cys Val Gln Thr Cys Val Asn Thr  
190 195 200

Tyr Gly Ser Phe Ile Cys Arg Cys Asp Pro Gly Tyr Glu Leu Glu Glu  
205 210 215

Asp Gly Ile His Cys Ser Asp Met Asp Glu Cys Ser Phe Ser Glu Phe  
220 225 230

Leu Cys Gln His Glu Cys Val Asn Gln Pro Gly Ser Tyr Phe Cys Ser

235                    240                    245

Cys Pro Pro Gly Tyr Val Leu Leu Asp Asp Asn Arg Ser Cys Gln Asp

250                    255                    260                    265

Ile Asn Glu Cys Glu His Arg Asn His Thr Cys Thr Ser Leu Gln Thr

270                    275                    280

Cys Tyr Asn Leu Gln Gly Gly Phe Lys Cys Ile Asp Pro Ile Ser Cys

285                    290                    295

Glu Glu Pro Tyr Leu Leu Ile Gly Glu Asn Arg Cys Met Cys Pro Ala

300                    305                    310

Glu His Thr Ser Cys Arg Asp Gln Pro Phe Thr Ile Leu Tyr Arg Asp

315                    320                    325

Met Asp Val Val Ser Gly Arg Ser Val Pro Ala Asp Ile Phe Gln Met

330                    335                    340                    345

Gln Ala Thr Thr Arg Tyr Pro Gly Ala Tyr Tyr Ile Phe Gln Ile Lys

350                    355                    360

Ser Gly Asn Glu Gly Arg Glu Phe Tyr Met Arg Gln Thr Gly Pro Ile

365                    370                    375

Ser Ala Thr Leu Val Met Thr Arg Pro Ile Lys Gly Pro Arg Asp Ile

380

385

390

Gln Leu Asp Leu Glu Met Ile Thr Val Asn Thr Val Ile Asn Phe Arg

395

400

405

Gly Ser Ser Val Ile Arg Leu Arg Ile Tyr Val Ser Gln Tyr Pro Phe

410

415

420

425

SEQ ID NO. : 2

Length : 1344 base pairs

Type : nucleic acid

Strandness : single

Topology : liner

Molecule type : cDNA to mRNA

Sequence Description

ATGCCAGGAT TAAAAAGGAT ACTCACTGTT ACCATCTTGG CACTCTGGCT TCCACATCCT 60

GGGAATGCAC AGCAGCAGTG CACAAACGGC TTTGACCTGG ACCGCCAGTC AGGACAGTGT 120

CTAGATATTG ATGAATGCCG GACCATCCCT GAGGCTTGTC GTGGGGACAT GATGTGTGTC 180

AACCAGAATG GCGGGTATTT GTGCATCCCT CGAACCAAACC CAGTGTATCG AGGGCCTTAC 240

TCAAATCCCT ACTCTACATC CTACTCAGGC CCATACCCAG CAGCGGCCCC ACCAGTACCA 300

GCTTCCAACCT ACCCCACGAT TTCAAGGCCT CTTGTCTGCC GCTTTGGGTA TCAGATGGAT 360

GAAGGCAACC AGTGTGTGGA TGTGGACGAG TGTGCAACAG ACTCACACCA GTGCAACCCT 420  
  
ACCCAGATCT GTATCAACAC TGAAGGAGGT TACACCTGCT CCTGCACCGA TGGGTACTGG 480  
  
CTTCTGGAAG GGCAGTGCT AGATATTGAT GAATGTCGCT ATGGTTACTG CCAGCAGCTC 540  
  
TGTGCAAATG TTCCAGGATC CTATTCTGT ACATGCAACC CTGGTTTCAC CCTCAACGAC 600  
  
GATGGAAGGT CTTGCCAAGA TGTGAACGAG TGCGAAACTG AGAATCCCTG TGTCAGACC 660  
  
TGTGTCAACA CCTATGGCTC TTTCATCTGC CGCTGTGACC CAGGATATGA ACTTGAGGAA 720  
  
GATGGCATTG ACTGCAGTGAT TATGGACGAG TGCGAGCTTCT CCGAGTTCCCT CTGTCAACAC 780  
  
GAGTGTGTGA ACCAGCCGGG CTCATACTTC TGCTCGTGCC CTCCAGGCTA CGTCCTGTTG 840  
  
GATGATAACC GAAGCTGCCA GGATATCAAT GAATGTGAGC ACCGAAACCA CACGTGTACC 900  
  
TCACTGCAGA CTTGCTACAA TCTACAAGGG GGCTTCAAAT GTATTGATCC CATCAGCTGT 960  
  
GAGGAGCCTT ATCTGCTGAT TGGTAAAAC CGCTGTATGT GTCCTGCTGA GCACACCAGC 1020  
  
TGCAGAGACC AGCCATTACAC CATCCTGTAT CGGGACATGG ATGTGGTGTGTC AGGACGCTCC 1080  
  
GTTCCTGCTG ACATCTTCCA GATGCAAGCA ACAACCCGAT ACCCTGGTGC CTATTACATT 1140  
  
TTCCAGATCA AATCTGGCAA CGAGGGTCGA GAGTTCTATA TGCGGCAAAC AGGGCCTATC 1200

AGTGCCACCC TGGTGATGAC ACGCCCCATC AAAGGGCCTC GGGACATCCA GCTGGACTTG 1260

GAGATGATCA CTGTCAACAC TGTCAATCAAC TTCAGAGGCA GCTCCGTGAT CCGACTGC GG 1320

ATATATGTGT CGCAGTATCC GTTC 1344

SEQ ID NO. : 3

Length : 2233 base pairs

Type : nucleic acid

Strandness : single

Topology : liner

Molecule type : cDNA to mRNA

Original source

Organism : Mus Musculus

Organelle : day13 mouse embryonic heart

Clone Name : mouse A55

Feature

Name/Key : CDS

Location : 75..1418

Identification method : S

Name/Key : sig peptide

Location : 75..143

Identification method : S

Name/Key : mat peptide

Location : 144..1418

Identification method : S

Sequence Description:

AATTCGGCAC	GAGCCCCAGT	CCCACCGCAG	AGCCTGCCTT	CCTCGCGTCG	CTTCTCCCTCC	60
CGCGCATCTT GGAT ATG CCA GGA TTA AAA AGG ATA CTC ACT GTT ACC ATC						110
Met Pro Gly Leu Lys Arg Ile Leu Thr Val Thr Ile						
-23	-20		-15			
TTG GCA CTC TGG CTT CCA CAT CCT GGG AAT GCA CAG CAG CAG TGC ACA						158
Leu Ala Leu Trp Leu Pro His Pro Gly Asn Ala Gln Gln Gln Cys Thr						
-10	-5	1		5		
AAC GGC TTT GAC CTG GAC CGC CAG TCA GGA CAG TGT CTA GAT ATT GAT						206
Asn Gly Phe Asp Leu Asp Arg Gln Ser Gly Gln Cys Leu Asp Ile Asp						
10	15		20			
GAA TGC CGG ACC ATC CCT GAG GCT TGT CGT GGG GAC ATG ATG TGT GTC						254
Glu Cys Arg Thr Ile Pro Glu Ala Cys Arg Gly Asp Met Met Cys Val						
25	30		35			
AAC CAG AAT GGC GGG TAT TTG TGC ATC CCT CGA ACC AAC CCA GTG TAT						302
Asn Gln Asn Gly Gly Tyr Leu Cys Ile Pro Arg Thr Asn Pro Val Tyr						
40	45		50			
CGA GGG CCT TAC TCA AAT CCC TAC TCT ACA TCC TAC TCA GGC CCA TAC						350
Arg Gly Pro Tyr Ser Asn Pro Tyr Ser Thr Ser Tyr Ser Gly Pro Tyr						
55	60		65			
CCA GCA GCG GCC CCA CCA GTA CCA GCT TCC AAC TAC CCC ACG ATT TCA						398
Pro Ala Ala Ala Pro Pro Val Pro Ala Ser Asn Tyr Pro Thr Ile Ser						

70	75	80	85
AGG CCT CTT GTC TGC CGC TTT GGG TAT CAG ATG GAT GAA GGC AAC CAG			446
Arg Pro Leu Val Cys Arg Phe Gly Tyr Gln Met Asp Glu Gly Asn Gln			
	90	95	100
TGT GTG GAT GTG GAC GAG TGT GCA ACA GAC TCA CAC CAG TGC AAC CCT			494
Cys Val Asp Val Asp Glu Cys Ala Thr Asp Ser His Gln Cys Asn Pro			
	105	110	115
ACC CAG ATC TGT ATC AAC ACT GAA GGA GGT TAC ACC TGC TCC TGC ACC			542
Thr Gln Ile Cys Ile Asn Thr Glu Gly Gly Tyr Thr Cys Ser Cys Thr			
	120	125	130
GAT GGG TAC TGG CTT CTG GAA GGG CAG TGC CTA GAT ATT GAT GAA TGT			590
Asp Gly Tyr Trp Leu Leu Glu Gly Gln Cys Leu Asp Ile Asp Glu Cys			
	135	140	145
CGC TAT GGT TAC TGC CAG CAG CTC TGT GCA AAT GTT CCA GGA TCC TAT			638
Arg Tyr Gly Tyr Cys Gln Gln Leu Cys Ala Asn Val Pro Gly Ser Tyr			
	150	155	160
			165
TCC TGT ACA TGC AAC CCT GGT TTC ACC CTC AAC GAC GAT GGA AGG TCT			686
Ser Cys Thr Cys Asn Pro Gly Phe Thr Leu Asn Asp Asp Gly Arg Ser			
	170	175	180
TGC CAA GAT GTG AAC GAG TGC GAA ACT GAG AAT CCC TGT GTT CAG ACC			734
Cys Gln Asp Val Asn Glu Cys Glu Thr Glu Asn Pro Cys Val Gln Thr			

185	190	195	
TGT GTC AAC ACC TAT GGC TCT TTC ATC TGC CGC TGT GAC CCA GGA TAT      782			
Cys Val Asn Thr Tyr Gly Ser Phe Ile Cys Arg Cys Asp Pro Gly Tyr			
200	205	210	
GAA CTT GAG GAA GAT GGC ATT CAC TGC AGT GAT ATG GAC GAG TGC AGC      830			
Glu Leu Glu Glu Asp Gly Ile His Cys Ser Asp Met Asp Glu Cys Ser			
215	220	225	
TTC TCC GAG TTC CTC TGT CAA CAC GAG TGT GTG AAC CAG CCG GGC TCA .    878			
Phe Ser Glu Phe Leu Cys Gln His Glu Cys Val Asn Gln Pro Gly Ser			
230	235	240	245
TAC TTC TGC TCG TGC CCT CCA GGC TAC GTC CTG TTG GAT GAT AAC CGA    926			
Tyr Phe Cys Ser Cys Pro Pro Gly Tyr Val Leu Leu Asp Asp Asn Arg			
250	255	260	
AGC TGC CAG GAT ATC AAT GAA TGT GAG CAC CGA AAC CAC ACG TGT ACC    974			
Ser Cys Gln Asp Ile Asn Glu Cys Glu His Arg Asn His Thr Cys Thr			
265	270	275	
TCA CTG CAG ACT TGC TAC AAT CTA CAA GGG GGC TTC AAA TGT ATT GAT    1022			
Ser Leu Gln Thr Cys Tyr Asn Leu Gln Gly Gly Phe Lys Cys Ile Asp			
280	285	290	
CCC ATC AGC TGT GAG GAG CCT TAT CTG CTG ATT GGT GAA AAC CGC TGT    1070			
Pro Ile Ser Cys Glu Glu Pro Tyr Leu Leu Ile Gly Glu Asn Arg Cys			

295	300	305	
ATG TGT CCT GCT GAG CAC ACC AGC TGC AGA GAC CAG CCA TTC ACC ATC      1118			
Met Cys Pro Ala Glu His Thr Ser Cys Arg Asp Gln Pro Phe Thr Ile			
310	315	320	325
CTG TAT CGG GAC ATG GAT GTG GTG TCA GGA CGC TCC GTT CCT GCT GAC      1166			
Leu Tyr Arg Asp Met Asp Val Val Ser Gly Arg Ser Val Pro Ala Asp			
330	335	340	
ATC TTC CAG ATG CAA GCA ACA ACC CGA TAC CCT GGT GCC TAT TAC ATT      1214			
Ile Phe Gln Met Gln Ala Thr Thr Arg Tyr Pro Gly Ala Tyr Tyr Ile			
345	350	355	
TTC CAG ATC AAA TCT GGC AAC GAG GGT CGA GAG TTC TAT ATG CGG CAA      1262			
Phe Gln Ile Lys Ser Gly Asn Glu Gly Arg Glu Phe Tyr Met Arg Gln			
360	365	370	
ACA GGG CCT ATC AGT GCC ACC CTG GTG ATG ACA CGC CCC ATC AAA GGG      1310			
Thr Gly Pro Ile Ser Ala Thr Leu Val Met Thr Arg Pro Ile Lys Gly			
375	380	385	
CCT CGG GAC ATC CAG CTG GAC TTG GAG ATG ATC ACT GTC AAC ACT GTC      1358			
Pro Arg Asp Ile Gln Leu Asp Leu Glu Met Ile Thr Val Asn Thr Val			
390	395	400	405
ATC AAC TTC AGA GGC AGC TCC GTG ATC CGA CTG CGG ATA TAT GTG TCG      1406			
Ile Asn Phe Arg Gly Ser Ser Val Ile Arg Leu Arg Ile Tyr Val Ser			

410

415

420

CAG TAT CCG TTC TGAGCCTCTG GCTAAGGCCT CTGACACTGC CTTTCACCAG 1458

Gln Tyr Pro Phe

425

CACCGAGGGA CGGGAGGAGA AAGGAAACCA GCAAGAACATGA GAGCGAGACA GACATTGCAC 1518

CTTCCCTGCT GAATATCTCC TGGGGGCATC AGCCTAGCAT CTTGACCCAT ATCTGTACTA 1578

TTGCAGATGG TCACCTTGAA GGACACCCCTG CCCTCAGTTC CTATGATGCA GTTATCCAAA 1638

AGTGTTCATC TTAGCCCCCTG ATATGAGGTT GCCAGTGACT CTTCAAAGCC TTCCATTAT 1698

TTCCATCGTT TTATAAAAAA GAAAATAGAT TAGATTGCT GGGGTATGAG TCCTCGAAGG 1758

TTCAAAAGAC TGAGTGGCTT GCTCTCACCT CTTCCCTCTCC TTCCCTCCATC TCTTGCTGCA 1818

TTGCTGCTTT GCAAAAGTCC TCATGGGCTC GTGGAAATG CTGGGAATAG CTAGTTGCT 1878

TCTTGCATGT TCTGAGAAGG CTATGGAAC ACACCACAGC AGGATCGAAG GTTTTATAG 1938

AGTCTATTTT AAAATCACAT CTGGTATTTT CAGCATAAAA GAAATTTAG TTGTCTTAA 1998

AATTTGTATG AGTGTAAAC CTTTCTTAT TCATTTGAG GCTTCTTAA GTGGTAGAAT 2058

TCCTTCCAAA GGCCTCAGAT ACATGTTATG TTCAGTCTT CCAACCTCAT CCTTCCTGC 2118

ATCTTAGCCC AGTTTTACG AAGACCCCTT AATCATGCTT TNTTAAGAGT TTTTACCAA 2178

CTGCGTTGGA AGACAGAGGT ATCCAGACTG ATTAAATAAT TGAAGAAAAA AAAAAA 2233

SEQ ID NO. : 4

Length : 423 amino acids

Type : amino acid

Topology : liner

Molecule type : protein

Sequence Description :

Gln Cys Thr Asn Gly Phe Asp Leu Asp Arg Gln Ser Gly Gln Cys Leu

1 5 10 15

Asp Ile Asp Glu Cys Arg Thr Ile Pro Glu Ala Cys Arg Gly Asp Met

20 25 30

Met Cys Val Asn Gln Asn Gly Gly Tyr Leu Cys Ile Pro Arg Thr Asn

35 40 45

Pro Val Tyr Arg Gly Pro Tyr Ser Asn Pro Tyr Ser Thr Ser Tyr Ser

50 55 60

Gly Pro Tyr Pro Ala Ala Pro Pro Val Pro Ala Ser Asn Tyr Pro

65 70 75 80

Thr Ile Ser Arg Pro Leu Val Cys Arg Phe Gly Tyr Gln Met Asp Glu

85 90 95

Gly Asn Gln Cys Val Asp Val Asp Glu Cys Ala Thr Asp Ser His Gln  
100 105 110

Cys Asn Pro Thr Gln Ile Cys Ile Asn Thr Glu Gly Gly Tyr Thr Cys  
115 120 125

Ser Cys Thr Asp Gly Tyr Trp Leu Leu Glu Gly Gln Cys Leu Asp Ile  
130 135 140

Asp Glu Cys Arg Tyr Gly Tyr Cys Gln Gln Leu Cys Ala Asn Val Pro  
145 150 155 160

Gly Ser Tyr Ser Cys Thr Cys Asn Pro Gly Phe Thr Leu Asn Asp Asp  
165 170 175

Gly Arg Ser Cys Gln Asp Val Asn Glu Cys Glu Thr Glu Asn Pro Cys  
180 185 190

Val Gln Thr Cys Val Asn Thr Tyr Gly Ser Phe Ile Cys Arg Cys Asp  
195 200 205

Pro Gly Tyr Glu Leu Glu Glu Asp Gly Ile His Cys Ser Asp Met Asp  
210 215 220

Glu Cys Ser Phe Ser Glu Phe Leu Cys Gln His Glu Cys Val Asn Gln  
225 230 235 240

Pro Gly Ser Tyr Phe Cys Ser Cys Pro Pro Gly Tyr Val Leu Leu Asp  
245 250 255

Asp Asn Arg Ser Cys Gln Asp Ile Asn Glu Cys Glu His Arg Asn His  
260 265 270

Thr Cys Thr Ser Leu Gln Thr Cys Tyr Asn Leu Gln Gly Gly Phe Lys  
275 280 285

Cys Ile Asp Pro Ile Ser Cys Glu Glu Pro Tyr Leu Leu Ile Gly Glu  
290 295 300

Asn Arg Cys Met Cys Pro Ala Glu His Thr Ser Cys Arg Asp Gln Pro  
305 310 315 320

Phe Thr Ile Leu Tyr Arg Asp Met Asp Val Val Ser Gly Arg Ser Val  
325 330 335

Pro Ala Asp Ile Phe Gln Met Gln Ala Thr Thr Arg Tyr Pro Gly Ala  
340 345 350

Tyr Tyr Ile Phe Gln Ile Lys Ser Gly Asn Glu Gly Arg Glu Phe Tyr  
355 360 365

Met Arg Gln Thr Gly Pro Ile Ser Ala Thr Leu Val Met Thr Arg Pro  
370 375 380

Ile Lys Gly Pro Arg Asp Ile Gln Leu Asp Leu Glu Met Ile Thr Val

385                   390                   395                   400

Asn Thr Val Ile Asn Phe Arg Gly Ser Ser Val Ile Arg Leu Arg Ile

405                   410                   415

Tyr Val Ser Gln Tyr Pro Phe

420

SEQ ID NO. : 5

Length : 1269 base pairs

Type : nucleic acid

Strandness : single

Topology : liner

Molecule type : cDNA to mRNA

Sequence Description:

CAGTGCACAA ACGGCTTG A CCTGGACCGC CAGTCAGGAC AGTGTCTAGA TATTGATGAA       60

TGCCGGACCA TCCCTGAGGC TTGTCTGGG GACATGATGT GTGTCAACCA GAATGGCGGG       120

TATTTGTGCA TCCCTCGAAC CAACCCAGTG TATCGAGGGC CTTACTCAA TCCCTACTCT       180

ACATCCTACT CAGGCCATA CCCAGCAGCG GCCCCACCA TACCAGCTTC CAACTACCCC       240

ACGATTTCAA GGCCTCTTGT CTGCCGCTTT GGGTATCAGA TGGATGAAGG CAACCAGTGT       300

GTGGATGTGG ACGAGTGTGC AACAGACTCA CACCAAGTGCA ACCCTACCCA GATCTGTATC       360

AACACTGAAG GAGGTTACAC CTGCTCCTGC ACCGATGGGT ACTGGCTTCT GGAAGGGCAG 420

TGCCTAGATA TTGATGAATG TCGCTATGGT TACTGCCAGC AGCTCTGTGC AAATGTTCCA 480

GGATCCTATT CCTGTACATG CAACCCTGGT TTCACCCCTCA ACGACGATGG AAGGTCTTGC 540

CAAGATGTGA ACGAGTGCAG AACTGAGAAT CCCTGTGTT AGACACTGTGT CAACACCTAT 600

GGCTCTTCAT TCTGCCGCTG TGACCCAGGA TATGAACCTTG AGGAAGATGG CATTCACTGC 660

AGTGATATGG ACGAGTGCAG CTTCTCCGAG TTCCTCTGTC AACACGAGTG TGTGAACCAG 720

CCGGGCTCAT ACTTCTGCTC GTGCCCTCCA GGCTACGTCC TGTTGGATGA TAACCGAAC 780

TGCCAGGATA TCAATGAATG TGAGCACCGA AACACACAGT GTACCTCACT GCAGACTTGC 840

TACAATCTAC AAGGGGGCTT CAAATGTATT GATCCCATCA GCTGTGAGGA GCCTTATCTG 900

CTGATTGGTG AAAACCGCTG TATGTGTCCT GCTGAGCACA CCAGCTGCAG AGACCAGCCA 960

TTCACCATCC TGTATCGGGA CATGGATGTG GTGTCAGGAC GCTCCGTTCC TGCTGACATC 1020

TTCCAGATGC AAGCAACAAAC CCGATACCCCT GGTGCCTATT ACATTTCCA GATCAAATCT 1080

GGCAACGAGG GTCGAGAGTT CTATATGCGG CAAACAGGGC CTATCAGTGC CACCCCTGGTG 1140

ATGACACGCC CCATCAAAGG GCCTCGGGAC ATCCAGCTGG ACTTGGAGAT GATCACTGTC 1200

AACACTGTCA TCAACTTCAG AGGCAGCTCC GTGATCCGAC TGGGGATATA TGTGTCGCAG 1260

TATCCGTTTC 1269

SEQ ID NO. : 6

Length : 461 amino acids

Type : amino acid

Topology : liner

Molecule type : protein

Sequence Description:

Met Gly Pro Arg Ser Phe Glu Pro Met His Ser Gly Leu Cys Arg Gln  
-36 -35 -30 -25

Arg Arg Met Ile Leu Thr Val Thr Ile Leu Ala Leu Trp Leu Pro His  
-20 -15 -10 -5

Pro Gly Asn Ala Gln Gln Cys Thr Asn Gly Phe Asp Leu Asp Arg  
1 5 10

Gln Ser Gly Gln Cys Leu Asp Ile Asp Glu Cys Arg Thr Ile Pro Glu  
15 20 25

Ala Cys Arg Gly Asp Met Met Cys Val Asn Gln Asn Gly Gly Tyr Leu  
30 35 40

Cys Ile Pro Arg Thr Asn Pro Val Tyr Arg Gly Pro Tyr Ser Asn Pro

45                    50                    55                    60  
Tyr Ser Thr Ser Tyr Ser Gly Pro Tyr Pro Ala Ala Ala Pro Pro Val  
65                    70                    75  
Pro Ala Ser Asn Tyr Pro Thr Ile Ser Arg Pro Leu Val Cys Arg Phe  
80                    85                    90  
Gly Tyr Gln Met Asp Glu Gly Asn Gln Cys Val Asp Val Asp Glu Cys  
95                    100                  105  
Ala Thr Asp Ser His Gln Cys Asn Pro Thr Gln Ile Cys Ile Asn Thr  
110                  115                  120  
Glu Gly Gly Tyr Thr Cys Ser Cys Thr Asp Gly Tyr Trp Leu Leu Glu  
125                  130                  135                  140  
Gly Gln Cys Leu Asp Ile Asp Glu Cys Arg Tyr Gly Tyr Cys Gln Gln  
145                  150                  155  
Leu Cys Ala Asn Val Pro Gly Ser Tyr Ser Cys Thr Cys Asn Pro Gly  
160                  165                  170  
Phe Thr Leu Asn Asp Asp Gly Arg Ser Cys Gln Asp Val Asn Glu Cys  
175                  180                  185  
Glu Thr Glu Asn Pro Cys Val Gln Thr Cys Val Asn Thr Tyr Gly Ser  
190                  195                  200

Phe Ile Cys Arg Cys Asp Pro Gly Tyr Glu Leu Glu Glu Asp Gly Ile  
205                210                215                220

His Cys Ser Asp Met Asp Glu Cys Ser Phe Ser Glu Phe Leu Cys Gln  
225                230                235

His Glu Cys Val Asn Gln Pro Gly Ser Tyr Phe Cys Ser Cys Pro Pro  
240                245                250

Gly Tyr Val Leu Leu Asp Asp Asn Arg Ser Cys Gln Asp Ile Asn Glu  
255                260                265

Cys Glu His Arg Asn His Thr Cys Thr Ser Leu Gln Thr Cys Tyr Asn  
270                275                280

Leu Gln Gly Gly Phe Lys Cys Ile Asp Pro Ile Ser Cys Glu Glu Pro  
285                290                295                300

Tyr Leu Leu Ile Gly Glu Asn Arg Cys Met Cys Pro Ala Glu His Thr  
305                310                315

Ser Cys Arg Asp Gln Pro Phe Thr Ile Leu Tyr Arg Asp Met Asp Val  
320                325                330

Val Ser Gly Arg Ser Val Pro Ala Asp Ile Phe Gln Met Gln Ala Thr  
335                340                345

Thr Arg Tyr Pro Gly Ala Tyr Tyr Ile Phe Gln Ile Lys Ser Gly Asn

350 355 360

Glu Gly Arg Glu Phe Tyr Met Arg Gln Thr Gly Pro Ile Ser Ala Thr

365 370 375 380

Leu Val Met Thr Arg Pro Ile Lys Gly Pro Arg Asp Ile Gln Leu Asp

385 390 395

Leu Glu Met Ile Thr Val Asn Thr Val Ile Asn Phe Arg Gly Ser Ser

400 405 410

Val Ile Arg Leu Arg Ile Tyr Val Ser Gln Tyr Pro Phe

415 420 425

SEQ ID NO. : 7

Length : 1383 base pairs

Type : nucleic acid

Strandness : single

Topology : liner

Molecule type : cDNA to mRNA

Sequence Description:

ATGGGACCTA GAAGTTTCGA GCCAATGCAC AGTGGACTCT GCAGACAGAG ACGCATGATA 60

CTCACTGTTA CCATCTTGGC ACTCTGGCTT CCACATCCTG GGAATGCACA GCAGCAGTGC 120

ACAAACGGCT TTGACCTGGA CCGCCAGTCA GGACAGTGTC TAGATATTGA TGAATGCCGG 180

ACCATCCCTG AGGCTTGTG TGGGGACATG ATGTGTGTCA ACCAGAATGG CGGGTATTTG 240

TGCATCCCTC GAACCAACCC AGTGTATCGA GGGCCTTACT CAAATCCCTA CTCTACATCC 300

TACTCAGGCC CATAACCCAGC AGCGGCCCCA CCAGTACCAG CTTCCAACTA CCCCACGATT 360

TCAAGGCCTC TTGTCTGCCG CTTTGGGTAT CAGATGGATG AAGGCAACCA GTGTGTGGAT 420

GTGGACGAGT GTGCAACAGA CTCACACCAAG TGCAACCCCTA CCCAGATCTG TATCAACACT 480

GAAGGAGGTT ACACCTGCTC CTGCACCGAT GGGTACTGGC TTCTGGAAGG GCAGTGCCTA 540

GATATTGATG AATGTCGCTA TGGTTACTGC CAGCAGCTCT GTGCAAATGT TCCAGGATCC 600

TATTCTGTGTA CATGCAACCC TGGTTTCACC CTCAACGACG ATGGAAGGTC TTGCCAAGAT 660

GTGAACGAGT GCGAAACTGA GAATCCCTGT GTTCAGACCT GTGTCAACAC CTATGGCTCT 720

TTCATCTGCC GCTGTGACCC AGGATATGAA CTTGAGGAAG ATGGCATTCA CTGCAGTGAT 780

ATGGACGAGT GCAGCTTCTC CGAGTTCCCTC TGTCAACACG AGTGTGTGAA CCAGCCGGGC 840

TCATACTTCT GCTCGTGCC C TCCAGGCTAC GTCCTGTTGG ATGATAACCG AAGCTGCCAG 900

GATATCAATG AATGTGAGCA CCGAAACAC ACgtgtacct CACTGCAGAC TTGCTACAAT 960

CTACAAGGGG GCTTCAAATG TATTGATCCC ATCAGCTGTG AGGAGCCTTA TCTGCTGATT 1020

GGTGAAAACC GCTGTATGTG TCCTGCTGAG CACACCAGCT GCAGAGACCA GCCATTCAAC 1080

ATCCTGTATC GGGACATGGA TGTGGTGTCA GGACGCTCCG TTCCTGCTGA CATCTTCCAG 1140

ATGCAAGCAA CAACCCGATA CCCTGGTGCC TATTACATTT TCCAGATCAA ATCTGGCAAC 1200

GAGGGTCGAG AGTTCTATAT GCGGCAAACA GGGCCTATCA GTGCCACCCT GGTGATGACA 1260

CGCCCCATCA AAGGGCCTCG GGACATCCAG CTGGACTTGG AGATGATCAC TGTCAACACT 1320

GTCATCAACT TCAGAGGCAG CTCCGTGATC CGACTGCGGA TATATGTGTC GCAGTATCCG 1380

TTC 1383

SEQ ID NO. : 8

Length : 2429 base pairs

Type : nucleic acid

Strandness : single

Topology : liner

Molecule type : cDNA to mRNA

Original source

Organism : Mus Musculus

Organelle : day13 embryonic heart

Clone Name : mouse A55b

Feature

Name/Key : CDS

Location : 232..1614

Identification method : S

Name/Key : sig peptide

Location : 232..339

Identification method : S

Name/ Key : mat peptide

Location : 340..1614

Identification method : S

Sequence Description:

CAGCATCTCG AGAGAGGCAG CAGACAAACCT CTCTAGGTCA TTTCTCTTC TTTTGAAA 60

GGGCAGCAAC GTTGTGCGCA GTTTATAAAA TATCACACTA CATGTTTTT AAATTTGGGA 120

GACTGCTGAC TACGGCACCA GCAATTGCTT TGCTGCGACG GCTGTGAGAC AAGCAGAAGT 180

CTCCGAACAC TTCTGTCTGC GTTTGCTCTA TGTGTGTGAT TTACAGAGGG A ATG GGA 237

Met Gly

-36 -35

CCT AGA AGT TTC GAG CCA ATG CAC AGT GGA CTC TGC AGA CAG AGA CGC 285

Pro Arg Ser Phe Glu Pro Met His Ser Gly Leu Cys Arg Gln Arg Arg

-30

-25

-20

ATG ATA CTC ACT GTT ACC ATC TTG GCA CTC TGG CTT CCA CAT CCT GGG 333

Met Ile Leu Thr Val Thr Ile Leu Ala Leu Trp Leu Pro His Pro Gly

-15

-10

-5

AAT GCA CAG CAG CAG TGC ACA AAC GGC TTT GAC CTG GAC CGC CAG TCA 381

Asn Ala Gln Gln Gln Cys Thr Asn Gly Phe Asp Leu Asp Arg Gln Ser

1 5 10

GGA CAG TGT CTA GAT ATT GAT GAA TGC CGG ACC ATC CCT GAG GCT TGT 429

Gly Gln Cys Leu Asp Ile Asp Glu Cys Arg Thr Ile Pro Glu Ala Cys

15 20 25 30

CGT GGG GAC ATG ATG TGT GTC AAC CAG AAT GGC GGG TAT TTG TGC ATC 477

Arg Gly Asp Met Met Cys Val Asn Gln Asn Gly Gly Tyr Leu Cys Ile

35 40 45

CCT CGA ACC AAC CCA GTG TAT CGA GGG CCT TAC TCA AAT CCC TAC TCT 525

Pro Arg Thr Asn Pro Val Tyr Arg Gly Pro Tyr Ser Asn Pro Tyr Ser

50 55 60

ACA TCC TAC TCA GGC CCA TAC CCA GCA GCG GCC CCA CCA GTA CCA GCT 573

Thr Ser Tyr Ser Gly Pro Tyr Pro Ala Ala Ala Pro Pro Val Pro Ala

65 70 75

TCC AAC TAC CCC ACG ATT TCA AGG CCT CTT GTC TGC CGC TTT GGG TAT 621

Ser Asn Tyr Pro Thr Ile Ser Arg Pro Leu Val Cys Arg Phe Gly Tyr

80 85 90

CAG ATG GAT GAA GGC AAC CAG TGT GTG GAT GTG GAC GAG TGT GCA ACA 669

Gln Met Asp Glu Gly Asn Gln Cys Val Asp Val Asp Glu Cys Ala Thr

95 100 105 110

GAC TCA CAC CAG TGC AAC CCT ACC CAG ATC TGT ATC AAC ACT GAA GGA 717

Asp Ser His Gln Cys Asn Pro Thr Gln Ile Cys Ile Asn Thr Glu Gly

115

120

125

GGT TAC ACC TGC TCC TGC ACC GAT GGG TAC TGG CTT CTG GAA GGG CAG 765

Gly Tyr Thr Cys Ser Cys Thr Asp Gly Tyr Trp Leu Leu Glu Gly Gln

130

135

140

TGC CTA GAT ATT GAT GAA TGT CGC TAT GGT TAC TGC CAG CAG CTC TGT 813

Cys Leu Asp Ile Asp Glu Cys Arg Tyr Gly Tyr Cys Gln Gln Leu Cys

145

150

155

GCA AAT GTT CCA GGA TCC TAT TCC TGT ACA TGC AAC CCT GGT TTC ACC 861

Ala Asn Val Pro Gly Ser Tyr Ser Cys Thr Cys Asn Pro Gly Phe Thr

160

165

170

CTC AAC GAC GAT GGA AGG TCT TGC CAA GAT GTG AAC GAG TGC GAA ACT 909

Leu Asn Asp Asp Gly Arg Ser Cys Gln Asp Val Asn Glu Cys Glu Thr

175

180

185

190

GAG AAT CCC TGT GTT CAG ACC TGT GTC AAC ACC TAT GGC TCT TTC ATC 957

Glu Asn Pro Cys Val Gln Thr Cys Val Asn Thr Tyr Gly Ser Phe Ile

195

200

205

TGC CGC TGT GAC CCA GGA TAT GAA CTT GAG GAA GAT GGC ATT CAC TGC 1005

Cys Arg Cys Asp Pro Gly Tyr Glu Leu Glu Asp Gly Ile His Cys

210

215

220

AGT GAT ATG GAC GAG TGC AGC TTC TCC GAG TTC CTC TGT CAA CAC GAG 1053

Ser Asp Met Asp Glu Cys Ser Phe Ser Glu Phe Leu Cys Gln His Glu

225

230

235

TGT GTG AAC CAG CCG GGC TCA TAC TTC TGC TCG TGC CCT CCA GGC TAC 1101

Cys Val Asn Gln Pro Gly Ser Tyr Phe Cys Ser Cys Pro Pro Gly Tyr

240

245

250

GTC CTG TTG GAT GAT AAC CGA AGC TGC CAG GAT ATC AAT GAA TGT GAG 1149

Val Leu Leu Asp Asp Asn Arg Ser Cys Gln Asp Ile Asn Glu Cys Glu

255

260

265

270

CAC CGA AAC CAC ACG TGT ACC TCA CTG CAG ACT TGC TAC AAT CTA CAA 1197

His Arg Asn His Thr Cys Thr Ser Leu Gln Thr Cys Tyr Asn Leu Gln

275

280

285

GGG GGC TTC AAA TGT ATT GAT CCC ATC AGC TGT GAG GAG CCT TAT CTG 1245

Gly Gly Phe Lys Cys Ile Asp Pro Ile Ser Cys Glu Glu Pro Tyr Leu

290

295

300

CTG ATT GGT GAA AAC CGC TGT ATG TGT CCT GCT GAG CAC ACC AGC TGC 1293

Leu Ile Gly Glu Asn Arg Cys Met Cys Pro Ala Glu His Thr Ser Cys

305

310

315

AGA GAC CAG CCA TTC ACC ATC CTG TAT CGG GAC ATG GAT GTG GTG TCA 1341

Arg Asp Gln Pro Phe Thr Ile Leu Tyr Arg Asp Met Asp Val Val Ser

320

325

330

GGA CGC TCC GTT CCT GCT GAC ATC TTC CAG ATG CAA GCA ACA ACC CGA 1389

Gly Arg Ser Val Pro Ala Asp Ile Phe Gln Met Gln Ala Thr Thr Arg

335                    340                    345                    350

TAC CCT GGT GCC TAT TAC ATT TTC CAG ATC AAA TCT GGC AAC GAG GGT    1437

Tyr Pro Gly Ala Tyr Tyr Ile Phe Gln Ile Lys Ser Gly Asn Glu Gly

355                    360                    365

CGA GAG TTC TAT ATG CGG CAA ACA GGG CCT ATC AGT GCC ACC CTG GTG    1485

Arg Glu Phe Tyr Met Arg Gln Thr Gly Pro Ile Ser Ala Thr Leu Val

370                    375                    380

ATG ACA CGC CCC ATC AAA GGG CCT CGG GAC ATC CAG CTG GAC TTG GAG    1533

Met Thr Arg Pro Ile Lys Gly Pro Arg Asp Ile Gln Leu Asp Leu Glu

385                    390                    395

ATG ATC ACT GTC AAC ACT GTC ATC AAC TTC AGA GGC AGC TCC GTG ATC    1581

Met Ile Thr Val Asn Thr Val Ile Asn Phe Arg Gly Ser Ser Val Ile

400                    405                    410

CGA CTG CGG ATA TAT GTG TCG CAG TAT CCG TTC TGAGCCTCTG GCTAAGGCCT 1634

Arg Leu Arg Ile Tyr Val Ser Gln Tyr Pro Phe

415                    420                    425

CTGACACTGC CTTTCACCAG CACCGAGGGA CGGGAGGAGA AAGGAAACCA GCAAGAACAT 1694

GAGCGAGACA GACATTGCAC CTTTCCTGCT GAATATCTCC TGGGGGCATC AGCCTAGCAT 1754

CTTGACCCAT ATCTGTACTA TTGCAGATGG TCACTCTGAA GGACACCCTG CCCTCAGTTC 1814

CTATGATGCA GTTATCCAAA AGTGTTCATC TTAGCCCCCTG ATATGAGGTT GCCAGTGACT 1874

CTTCAAAGCC TTCCATTAT TTCCATCGTT TTATAAAAAA GAAAATAGAT TAGATTTGCT 1934

GGGGTATGAG TCCTCGAAGG TTCAAAAGAC TGAGTGGCTT GCTCTCACCT CTTCTCTCC 1994

TTCCTCCATC TCTTGCTGCA TTGCTGCTTT GCAAAAGTCC TCATGGGCTC GTGGAAATG 2054

CTGGGAATAG CTAGTTGCT TCTTGCATGT TCTGAGAAGG CTATGGAAC ACACCACAGC 2114

AGGATCGAAG GTTTTATAG AGTCTATTAA AAAATCACAT CTGGTATTAA CAGCATAAAA 2174

GAAATTTAG TTGTCTTAA AATTTGTATG AGTGTAAAC CTTTCTTAT TCATTTGAG 2234

GCTTCTTAA GTGGTAGAAT TCCTTCCAAA GGCCTCAGAT ACATGTTATG TTCAGTCTTT 2294

CCAACCTCAT CCTTCCTGC ATCTTAGCCC AGTTTACG AAGACCCCTT AATCATGCTT 2354

TNTTAAGAGT TTTACCCAA CTGCGTTGGA AGACAGAGGT ATCCAGACTG ATTAAATAAT 2414

TGAAGAAAAA AAAAAA 2429

SEQ ID NO. : 9

Length : 423 amino acids

Type : amino acid

Topology : liner

Molecule type : protein

Sequence Description :

Gln Cys Thr Asn Gly Phe Asp Leu Asp Arg Gln Ser Gly Gln Cys Leu

1 5 10 15

Asp Ile Asp Glu Cys Arg Thr Ile Pro Glu Ala Cys Arg Gly Asp Met

20 25 30

Met Cys Val Asn Gln Asn Gly Gly Tyr Leu Cys Ile Pro Arg Thr Asn

35 40 45

Pro Val Tyr Arg Gly Pro Tyr Ser Asn Pro Tyr Ser Thr Ser Tyr Ser

50 55 60

Gly Pro Tyr Pro Ala Ala Ala Pro Pro Val Pro Ala Ser Asn Tyr Pro

65 70 75 80

Thr Ile Ser Arg Pro Leu Val Cys Arg Phe Gly Tyr Gln Met Asp Glu

85 90 95

Gly Asn Gln Cys Val Asp Val Asp Glu Cys Ala Thr Asp Ser His Gln

100 105 110

Cys Asn Pro Thr Gln Ile Cys Ile Asn Thr Glu Gly Gly Tyr Thr Cys

115 120 125

Ser Cys Thr Asp Gly Tyr Trp Leu Leu Glu Gly Gln Cys Leu Asp Ile

130 135 140

Asp Glu Cys Arg Tyr Gly Tyr Cys Gln Gln Leu Cys Ala Asn Val Pro  
145 150 155 160

Gly Ser Tyr Ser Cys Thr Cys Asn Pro Gly Phe Thr Leu Asn Asp Asp  
165 170 175

Gly Arg Ser Cys Gln Asp Val Asn Glu Cys Glu Thr Glu Asn Pro Cys  
180 185 190

Val Gln Thr Cys Val Asn Thr Tyr Gly Ser Phe Ile Cys Arg Cys Asp  
195 200 205

Pro Gly Tyr Glu Leu Glu Glu Asp Gly Ile His Cys Ser Asp Met Asp  
210 215 220

Glu Cys Ser Phe Ser Glu Phe Leu Cys Gln His Glu Cys Val Asn Gln  
225 230 235 240

Pro Gly Ser Tyr Phe Cys Ser Cys Pro Pro Gly Tyr Val Leu Leu Asp  
245 250 255

Asp Asn Arg Ser Cys Gln Asp Ile Asn Glu Cys Glu His Arg Asn His  
260 265 270

Thr Cys Thr Ser Leu Gln Thr Cys Tyr Asn Leu Gln Gly Gly Phe Lys  
275 280 285

Cys Ile Asp Pro Ile Ser Cys Glu Glu Pro Tyr Leu Leu Ile Gly Glu

290

295

300

Asn Arg Cys Met Cys Pro Ala Glu His Thr Ser Cys Arg Asp Gln Pro

305

310

315

320

Phe Thr Ile Leu Tyr Arg Asp Met Asp Val Val Ser Gly Arg Ser Val

325

330

335

Pro Ala Asp Ile Phe Gln Met Gln Ala Thr Thr Arg Tyr Pro Gly Ala

340

345

350

Tyr Tyr Ile Phe Gln Ile Lys Ser Gly Asn Glu Gly Arg Glu Phe Tyr

355

360

365

Met Arg Gln Thr Gly Pro Ile Ser Ala Thr Leu Val Met Thr Arg Pro

370

375

380

Ile Lys Gly Pro Arg Asp Ile Gln Leu Asp Leu Glu Met Ile Thr Val

385

390

395

400

Asn Thr Val Ile Asn Phe Arg Gly Ser Ser Val Ile Arg Leu Arg Ile

405

410

415

Tyr Val Ser Gln Tyr Pro Phe

420

SEQ ID NO. : 10

Length : 1269 base pairs

Type : nucleic acid

Strandness : single

Topology : liner

Molecule type : cDNA to mRNA

Sequence Description :

CAGTGCACAA	ACGGCTTG	A	CCTGGACCGC	CAGTCAGGAC	AGTGTCTAGA	TATTGATGAA	60	
TGCCGGACCA	TCCCTGAGGC	TTGTCGTGGG	GACATGATGT	GTGTCAACCA	GAATGGCGGG	120		
TATTTGTGCA	TCCCTCGAAC	CAACCCAGTG	TATCGAGGGC	CTTACTCAAA	TCCCTACTCT	180		
ACATCCTACT	CAGGCCATA	CCCAGCAGCG	GCCCCACCAG	TACCAGCTTC	CAAAC	TAACCCCC	240	
ACGATTCAA	GGCCTCTTGT	CTGCCGCTTT	GGGTATCAGA	TGGATGAAGG	CAACCAGTGT	300		
GTGGATGTGG	ACGAGTGTGC	AACAGACTCA	CACCA	GTGCA	ACCCTACCCA	GATCTGTATC	360	
AACACTGAAG	GAGGTTACAC	CTGCTCCTGC	ACCGATGGGT	ACTGGCTTCT	GGAAGGGCAG	420		
TGCCTAGATA	TTGATGAATG	TCGCTATGGT	TACTGCCAGC	AGCTCTGTGC	AAATGTTCCA	480		
GGATCCTATT	CCTGTACATG	CAACCCTGGT	TTCACCCCTCA	ACGACGATGG	AAGGTCTTGC	540		
CAAGATGTGA	ACGAGTGC	GA	AACTGAGAAT	CCCTGTGTT	AGACCTGTGT	CAACACCTAT	600	
GGCTCTTCA	TCTGCCGCTG	TGACCCAGGA	TATGAAC	TTG	AGGAAGATGG	CATTC	ACTGC	660

AGTGATATGG ACGAGTGCAG CTTCTCCGAG TTCCTCTGTC AACACGAGTG TGTGAACCAG 720  
  
CCGGGCTCAT ACTTCTGCTC GTGCCCTCCA GGCTACGTCC TGTTGGATGA TAACCGAAGC 780  
  
TGCCAGGATA TCAATGAATG TGAGCACCGA AACACACGT GTACCTCACT GCAGACTTGC 840  
  
TACAATCTAC AAGGGGGCTT CAAATGTATT GATCCCATCA GCTGTGAGGA GCCTTATCTG 900  
  
CTGATTGGTG AAAACCGCTG TATGTGTCCT GCTGAGCACA CCAGCTGCAG AGACCAGCCA 960  
  
TTCACCATCC TGTATCGGGA CATGGATGTG GTGTCAGGAC GCTCCGTTCC TGCTGACATC 1020  
  
TTCCAGATGC AAGCAACAAAC CCGATAACCT GGTGCCTATT ACATTTCCA GATCAAATCT 1080  
  
GGCAACGAGG GTCGAGAGTT CTATATGCGG CAAACAGGGC CTATCAGTGC CACCCCTGGTG 1140  
  
ATGACACGCC CCATCAAAGG GCCTCGGGAC ATCCAGCTGG ACTTGGAGAT GATCACTGTC 1200  
  
AACACTGTCA TCAACTTCAG AGGCAGCTCC GTGATCCGAC TGCAGATATA TGTGTCGCAG 1260  
  
TATCCGTTTC 1269

SEQ ID NO. : 11

Length : 448 amino acids

Type : amino acid

Topology : liner

Molecule type : protein

**Sequence Description :**

Met Pro Gly Ile Lys Arg Ile Leu Thr Val Thr Ile Leu Ala Leu Cys

-23 -20 -15 -10

Leu Pro Ser Pro Gly Asn Ala Gln Ala Gln Cys Thr Asn Gly Phe Asp

-5 1 5

Leu Asp Arg Gln Ser Gly Gln Cys Leu Asp Ile Asp Glu Cys Arg Thr

10 15 20 25

Ile Pro Glu Ala Cys Arg Gly Asp Met Met Cys Val Asn Gln Asn Gly

30 35 40

Gly Tyr Leu Cys Ile Pro Arg Thr Asn Pro Val Tyr Arg Gly Pro Tyr

45 50 55

Ser Asn Pro Tyr Ser Thr Pro Tyr Ser Gly Pro Tyr Pro Ala Ala Ala

60 65 70

Pro Pro Leu Ser Ala Pro Asn Tyr Pro Thr Ile Ser Arg Pro Leu Ile

75 80 85

Cys Arg Phe Gly Tyr Gln Met Asp Glu Ser Asn Gln Cys Val Asp Val

90 95 100 105

Asp Glu Cys Ala Thr Asp Ser His Gln Cys Asn Pro Thr Gln Ile Cys

110 115 120

Ile Asn Thr Glu Gly Gly Tyr Thr Cys Ser Cys Thr Asp Gly Tyr Trp

125                    130                    135

Leu Leu Glu Gly Gln Cys Leu Asp Ile Asp Glu Cys Arg Tyr Gly Tyr

140                    145                    150

Cys Gln Gln Leu Cys Ala Asn Val Pro Gly Ser Tyr Ser Cys Thr Cys

155                    160                    165

Asn Pro Gly Phe Thr Leu Asn Glu Asp Gly Arg Ser Cys Gln Asp Val

170                    175                    180                    185

Asn Glu Cys Ala Thr Glu Asn Pro Cys Val Gln Thr Cys Val Asn Thr

190                    195                    200

Tyr Gly Ser Phe Ile Cys Arg Cys Asp Pro Gly Tyr Glu Leu Glu Glu

205                    210                    215

Asp Gly Val His Cys Ser Asp Met Asp Glu Cys Ser Phe Ser Glu Phe

220                    225                    230

Leu Cys Gln His Glu Cys Val Asn Gln Pro Gly Thr Tyr Phe Cys Ser

235                    240                    245

Cys Pro Pro Gly Tyr Ile Leu Leu Asp Asp Asn Arg Ser Cys Gln Asp

250                    255                    260                    265

Ile Asn Glu Cys Glu His Arg Asn His Thr Cys Asn Leu Gln Gln Thr

270

275

280

Cys Tyr Asn Leu Gln Gly Gly Phe Lys Cys Ile Asp Pro Ile Arg Cys

285

290

295

Glu Glu Pro Tyr Leu Arg Ile Ser Asp Asn Arg Cys Met Cys Pro Ala

300

305

310

Glu Asn Pro Gly Cys Arg Asp Gln Pro Phe Thr Ile Leu Tyr Arg Asp

315

320

325

Met Asp Val Val Ser Gly Arg Ser Val Pro Ala Asp Ile Phe Gln Met

330

335

340

345

Gln Ala Thr Thr Arg Tyr Pro Gly Ala Tyr Tyr Ile Phe Gln Ile Lys

350

355

360

Ser Gly Asn Glu Gly Arg Glu Phe Tyr Met Arg Gln Thr Gly Pro Ile

365

370

375

Ser Ala Thr Leu Val Met Thr Arg Pro Ile Lys Gly Pro Arg Glu Ile

380

385

390

Gln Leu Asp Leu Glu Met Ile Thr Val Asn Thr Val Ile Asn Phe Arg

395

400

405

Gly Ser Ser Val Ile Arg Leu Arg Ile Tyr Val Ser Gln Tyr Pro Phe

410

415

420

425

SEQ ID NO. : 12

Length : 1344 base pairs

Type : nucleic acid

Strandness : single

Topology : liner

Molecule type : cDNA to mRNA

Sequence Description :

ATGCCAGGAA TAAAAAGGAT ACTCACTGTT ACCATTCTGG CTCTCTGTCT TCCAAGCCCT 60  
GGGAATGCAC AGGCACAGTG CACGAATGGC TTTGACCTGG ATGCCAGTC AGGACAGTGT 120  
TTAGATATTG ATGAATGCCG AACCATCCCC GAGGCCTGCC GAGGAGACAT GATGTGTGTT 180  
AACCAAAATG GCGGGTATTT ATGCATTCCC CGGACAAACC CTGTGTATCG AGGGCCCTAC 240  
TCGAACCCCT ACTCGACCCC CTACTCAGGT CCGTACCCAG CAGCTGCCCT ACCACTCTCA 300  
GCTCCAAACT ATCCCACGAT CTCCAGGCCT CTTATATGCC GCTTTGGATA CCAGATGGAT 360  
GAAAGCAACC AATGTGTGGA TGTGGACGAG TGTGCAACAG ATTCCCACCA GTGCAACCCC 420  
ACCCAGATCT GCATCAATAC TGAAGGCGGG TACACCTGCT CCTGCACCGA CGGATATTGG 480  
CTTCTGGAAG GCCAGTGCTT AGACATTGAT GAATGTCGCT ATGGTTACTG CCAGCAGCTC 540

TGTGCGAATG TTCCTGGATC CTATTCTTGT ACATGCAACC CTGGTTTAC CCTCAATGAG 600  
  
GATGGAAGGT CTTGCCAAGA TGTGAACGAG TGTGCCACCG AGAACCCCTG CGTGCAAACC 660  
  
TGCCTCAACA CCTACGGCTC TTTCATCTGC CGCTGTGACC CAGGATATGA ACTTGAGGAA 720  
  
GATGGCGTTC ATTGCAGTGA TATGGACGAG TGCAGCTTCT CTGAGTTCCCT CTGCCAACAT 780  
  
GAGTGTGTGA ACCAGCCCCG CACATACTTC TGCTCCTGCC CTCCAGGCTA CATCCTGCTG 840  
  
GATGACAACC GAAGCTGCCA AGACATCAAC GAATGTGAGC ACAGGAACCA CACGTGCAAC 900  
  
CTGCAGCAGA CGTGCTACAA TTTACAAGGG GGCTTCAAAT GCATCGACCC CATCCGCTGT 960  
  
GAGGAGCCTT ATCTGAGGAT CAGTGATAAC CGCTGTATGT GTCCTGCTGA GAACCCTGGC 1020  
  
TGCAGAGACC AGCCCTTAC CATCTTGTAC CGGGACATGG ACGTGGTGTGTC AGGACGCTCC 1080  
  
GTTCCCGCTG ACATCTTCCA AATGCAAGCC ACGACCCGCT ACCCTGGGGC CTATTACATT 1140  
  
TTCCAGATCA AATCTGGAA TGAGGGCAGA GAATTTACA TGC GGCAAAC GGGCCCCATC 1200  
  
AGTGCCACCC TGGTGATGAC ACGCCCCATC AAAGGGCCCC GGGAAATCCA GCTGGACTTG 1260  
  
GAAATGATCA CTGTCAACAC TGTCATCAAC TTCAGAGGCA GCTCCGTGAT CCGACTGCGG 1320  
  
ATATATGTGT CGCAGTACCC ATTC 1344

SEQ ID NO. : 13

Length : 2328 base pairs

Type : nucleic acid

Strandness : single

Topology : liner

Molecule type : cDNA to mRNA

Original source

Organism : Homo Sapiens

Organelle : human brain

Clone Name : human A55

Feature

Name/Key : CDS

Location : 169..1512

Identification method : S

Name/Key : sig peptide

Location : 169..237

Identification method : S

Name/ Key : mat peptide

Location : 238..1512

Identification method : S

Sequence Description :

GACCCGGCGC TCTCCCCGTG TCCTCTCCAC GACTCGCTCG GCCCCTCTGG AATAAAACAC 60

CCCGCGAGCCC CGAGGGCCCA GAGGAGGCCG ACGTGCCCCGA GCTCCTCCGG GGGTCCCGCC 120

CGCGAGCTTT CTTCTCGCCT TCGCATCTCC TCCTCGCGCG TCTTGGAC ATG CCA GGA 177

Met Pro Gly

ATA AAA AGG ATA CTC ACT GTT ACC ATT CTG GCT CTC TGT CTT CCA AGC	225		
Ile Lys Arg Ile Leu Thr Val Thr Ile Leu Ala Leu Cys Leu Pro Ser			
-20	-15	-10	-5
CCT GGG AAT GCA CAG GCA CAG TGC ACG AAT GGC TTT GAC CTG GAT CGC	273		
Pro Gly Asn Ala Gln Ala Gln Cys Thr Asn Gly Phe Asp Leu Asp Arg			
1	5	10	
CAG TCA GGA CAG TGT TTA GAT ATT GAT GAA TGC CGA ACC ATC CCC GAG	321		
Gln Ser Gly Gln Cys Leu Asp Ile Asp Glu Cys Arg Thr Ile Pro Glu			
15	20	25	
GCC TGC CGA GGA GAC ATG ATG TGT GTT AAC CAA AAT GGC GGG TAT TTA	369		
Ala Cys Arg Gly Asp Met Met Cys Val Asn Gln Asn Gly Gly Tyr Leu			
30	35	40	
TGC ATT CCC CGG ACA AAC CCT GTG TAT CGA GGG CCC TAC TCG AAC CCC	417		
Cys Ile Pro Arg Thr Asn Pro Val Tyr Arg Gly Pro Tyr Ser Asn Pro			
45	50	55	60
TAC TCG ACC CCC TAC TCA GGT CCG TAC CCA GCA GCT GCC CCA CCA CTC	465		
Tyr Ser Thr Pro Tyr Ser Gly Pro Tyr Pro Ala Ala Ala Pro Pro Leu			
65	70	75	
TCA GCT CCA AAC TAT CCC ACG ATC TCC AGG CCT CTT ATA TGC CGC TTT	513		
Ser Ala Pro Asn Tyr Pro Thr Ile Ser Arg Pro Leu Ile Cys Arg Phe			

80	85	90	
GGA TAC CAG ATG GAT GAA AGC AAC CAA TGT GTG GAT GTG GAC GAG TGT			
Gly Tyr Gln Met Asp Glu Ser Asn Gln Cys Val Asp Val Asp Glu Cys			
95	100	105	
GCA ACA GAT TCC CAC CAG TGC AAC CCC ACC CAG ATC TGC ATC AAT ACT			
Ala Thr Asp Ser His Gln Cys Asn Pro Thr Gln Ile Cys Ile Asn Thr			
110	115	120	
GAA GGC GGG TAC ACC TGC TCC TGC ACC GAC GGA TAT TGG CTT CTG GAA			
Glu Gly Gly Tyr Thr Cys Ser Cys Thr Asp Gly Tyr Trp Leu Leu Glu			
125	130	135	140
GGC CAG TGC TTA GAC ATT GAT GAA TGT CGC TAT GGT TAC TGC CAG CAG			
Gly Gln Cys Leu Asp Ile Asp Glu Cys Arg Tyr Gly Tyr Cys Gln Gln			
145	150	155	
CTC TGT GCG AAT GTT CCT GGA TCC TAT TCT TGT ACA TGC AAC CCT GGT			
Leu Cys Ala Asn Val Pro Gly Ser Tyr Ser Cys Thr Cys Asn Pro Gly			
160	165	170	
TTT ACC CTC AAT GAG GAT GGA AGG TCT TGC CAA GAT GTG AAC GAG TGT			
Phe Thr Leu Asn Glu Asp Gly Arg Ser Cys Gln Asp Val Asn Glu Cys			
175	180	185	
GCC ACC GAG AAC CCC TGC GTG CAA ACC TGC GTC AAC ACC TAC GGC TCT			
Ala Thr Glu Asn Pro Cys Val Gln Thr Cys Val Asn Thr Tyr Gly Ser			

190	195	200	
TTC ATC TGC CGC TGT GAC CCA GGA TAT GAA CTT GAG GAA GAT GGC GTT      897			
Phe Ile Cys Arg Cys Asp Pro Gly Tyr Glu Leu Glu Glu Asp Gly Val			
205	210	215	220
CAT TGC AGT GAT ATG GAC GAG TGC AGC TTC TCT GAG TTC CTC TGC CAA      945			
His Cys Ser Asp Met Asp Glu Cys Ser Phe Ser Glu Phe Leu Cys Gln			
225	230	235	
CAT GAG TGT GTG AAC CAG CCC GGC ACA TAC TTC TGC TCC TGC CCT CCA      993			
His Glu Cys Val Asn Gln Pro Gly Thr Tyr Phe Cys Ser Cys Pro Pro			
240	245	250	
GGC TAC ATC CTG CTG GAT GAC AAC CGA AGC TGC CAA GAC ATC AAC GAA      1041			
Gly Tyr Ile Leu Leu Asp Asp Asn Arg Ser Cys Gln Asp Ile Asn Glu			
255	260	265	
TGT GAG CAC AGG AAC CAC ACG TGC AAC CTG CAG CAG ACG TGC TAC AAT      1089			
Cys Glu His Arg Asn His Thr Cys Asn Leu Gln Gln Thr Cys Tyr Asn			
270	275	280	
TTA CAA GGG GGC TTC AAA TGC ATC GAC CCC ATC CGC TGT GAG GAG CCT      1137			
Leu Gln Gly Gly Phe Lys Cys Ile Asp Pro Ile Arg Cys Glu Glu Pro			
285	290	295	300
TAT CTG AGG ATC AGT GAT AAC CGC TGT ATG TGT CCT GCT GAG AAC CCT      1185			
Tyr Leu Arg Ile Ser Asp Asn Arg Cys Met Cys Pro Ala Glu Asn Pro			

305	310	315	
GGC TGC AGA GAC CAG CCC TTT ACC ATC TTG TAC CGG GAC ATG GAC GTG      1233			
Gly Cys Arg Asp Gln Pro Phe Thr Ile Leu Tyr Arg Asp Met Asp Val			
320	325	330	
GTG TCA GGA CGC TCC GTT CCC GCT GAC ATC TTC CAA ATG CAA GCC ACG      1281			
Val Ser Gly Arg Ser Val Pro Ala Asp Ile Phe Gln Met Gln Ala Thr			
335	340	345	
ACC CGC TAC CCT GGG GCC TAT TAC ATT TTC CAG ATC AAA TCT GGG AAT      1329			
Thr Arg Tyr Pro Gly Ala Tyr Tyr Ile Phe Gln Ile Lys Ser Gly Asn			
350	355	360	
GAG GGC AGA GAA TTT TAC ATG CGG CAA ACG GGC CCC ATC AGT GCC ACC      1377			
Glu Gly Arg Glu Phe Tyr Met Arg Gln Thr Gly Pro Ile Ser Ala Thr			
365	370	375	380
CTG GTG ATG ACA CGC CCC ATC AAA GGG CCC CGG GAA ATC CAG CTG GAC      1425			
Leu Val Met Thr Arg Pro Ile Lys Gly Pro Arg Glu Ile Gln Leu Asp			
385	390	395	
TTG GAA ATG ATC ACT GTC AAC ACT GTC ATC AAC TTC AGA GGC AGC TCC      1473			
Leu Glu Met Ile Thr Val Asn Thr Val Ile Asn Phe Arg Gly Ser Ser			
400	405	410	
GTG ATC CGA CTG CGG ATA TAT GTG TCG CAG TAC CCA TTC TGAGCCTCGG      1522			
Val Ile Arg Leu Arg Ile Tyr Val Ser Gln Tyr Pro Phe			

415

420

425

GCTGGAGCCT CCGACGCTGC CTCTCATTGG CACCAAGGGA CAGGAGAAGA GAGGAAATAA 1582

CAGAGAGAAT GAGAGCGACA CAGACGTTAG GCATTCCTG CTGAACGTTT CCCCCGAAGAG 1642

TCAGCCCCGA CTTCCTGACT CTCACCTGTA CTATTGCAGA CCTGTCACCC TGCAGGACTT 1702

GCCACCCCCA GTTCCTATGA TACAGTTATC AAAAAGTATT ATCATTGCTC CCCTGATAGA 1762

AGATTGTTGG TGAATTTCAG AGGCCTTCAG TTTATTTCCA CTATTTCAA AGAAAATAGA 1822

TTAGGTTTGC GGGGGTCTGA GTCTATGTTC AAAGACTGTG AACAGCTTGC TGTCACTTCT 1882

TCACCTCTTC CACTCCTTCT CTCACTGTGT TACTGCTTG CAAAGACCCG GGAGCTGGCG 1942

GGGAACCCTG GGAGTAGCTA GTTGCTTT TGCGTACACA GAGAAGGCTA TGTAAACAAA 2002

CCACAGCAGG ATCGAAGGGT TTTAGAGAA TGTGTTCAA AACCATGCCT GGTATTTCA 2062

ACCATAAAAG AAGTTTCAGT TGTCTTAAA TTTGTATAAC GGTTTAATTG TGTCTTGTTC 2122

ATTTTGAGTA TTTTAAAAAA ATATGTCGTA GAATTCCCTTC GAAAGGCCTT CAGACACATG 2182

CTATGTTCTG TCTTCCAAA CCCAGTCTCC TCTCCATTAGCAGTGT TTTCTTGAG 2242

GACCCCTTAA TCTTGCTTTC TTTAGAATTG TTACCCAATT GGATTGGAAT GCAGAGGTCT 2302

CCAAACTGAT TAAATATTTG AAGAGA

2328

SEQ ID NO. : 14

Length : 423 amino acids

Type : amino acid

Topology : liner

Molecule type : protein

Sequence Description :

Gln Cys Thr Asn Gly Phe Asp Leu Asp Arg Gln Ser Gly Gln Cys Leu

1 5 10 15

Asp Ile Asp Glu Cys Arg Thr Ile Pro Glu Ala Cys Arg Gly Asp Met

20 25 30

Met Cys Val Asn Gln Asn Gly Gly Tyr Leu Cys Ile Pro Arg Thr Asn

35 40 45

Pro Val Tyr Arg Gly Pro Tyr Ser Asn Pro Tyr Ser Thr Pro Tyr Ser

50 55 60

Gly Pro Tyr Pro Ala Ala Ala Pro Pro Leu Ser Ala Pro Asn Tyr Pro

65 70 75 80

Thr Ile Ser Arg Pro Leu Ile Cys Arg Phe Gly Tyr Gln Met Asp Glu

85 90 95

Ser Asn Gln Cys Val Asp Val Asp Glu Cys Ala Thr Asp Ser His Gln

100 105 110

Cys Asn Pro Thr Gln Ile Cys Ile Asn Thr Glu Gly Gly Tyr Thr Cys

115 120 125

Ser Cys Thr Asp Gly Tyr Trp Leu Leu Glu Gly Gln Cys Leu Asp Ile

130 135 140

Asp Glu Cys Arg Tyr Gly Tyr Cys Gln Gln Leu Cys Ala Asn Val Pro

145 150 155 160

Gly Ser Tyr Ser Cys Thr Cys Asn Pro Gly Phe Thr Leu Asn Glu Asp

165 170 175

Gly Arg Ser Cys Gln Asp Val Asn Glu Cys Ala Thr Glu Asn Pro Cys

180 185 190

Val Gln Thr Cys Val Asn Thr Tyr Gly Ser Phe Ile Cys Arg Cys Asp

195 200 205

Pro Gly Tyr Glu Leu Glu Glu Asp Gly Val His Cys Ser Asp Met Asp

210 215 220

Glu Cys Ser Phe Ser Glu Phe Leu Cys Gln His Glu Cys Val Asn Gln

225 230 235 240

Pro Gly Thr Tyr Phe Cys Ser Cys Pro Pro Gly Tyr Ile Leu Leu Asp

245 250 255

Asp Asn Arg Ser Cys Gln Asp Ile Asn Glu Cys Glu His Arg Asn His

260

265

270

Thr Cys Asn Leu Gln Gln Thr Cys Tyr Asn Leu Gln Gly Gly Phe Lys

275

280

285

Cys Ile Asp Pro Ile Arg Cys Glu Glu Pro Tyr Leu Arg Ile Ser Asp

290

295

300

Asn Arg Cys Met Cys Pro Ala Glu Asn Pro Gly Cys Arg Asp Gln Pro

305

310

315

320

Phe Thr Ile Leu Tyr Arg Asp Met Asp Val Val Ser Gly Arg Ser Val

325

330

335

Pro Ala Asp Ile Phe Gln Met Gln Ala Thr Thr Arg Tyr Pro Gly Ala

340

345

350

Tyr Tyr Ile Phe Gln Ile Lys Ser Gly Asn Glu Gly Arg Glu Phe Tyr

355

360

365

Met Arg Gln Thr Gly Pro Ile Ser Ala Thr Leu Val Met Thr Arg Pro

370

375

380

Ile Lys Gly Pro Arg Glu Ile Gln Leu Asp Leu Glu Met Ile Thr Val

385

390

395

400

Asn Thr Val Ile Asn Phe Arg Gly Ser Ser Val Ile Arg Leu Arg Ile

405

410

415

Tyr Val Ser Gln Tyr Pro Phe

420

SEQ ID NO. : 15

Length : 1269 base pairs

Type : nucleic acid

Strandness : single

Topology : liner

Molecule type : cDNA to mRNA

Sequence Description :

CAGTGCACGA ATGGCTTG A CCTGGATCGC CAGTCAGGAC AGTGTAGA TATTGATGAA 60

TGCCGAACCA TCCCCGAGGC CTGCCGAGGA GACATGATGT GTGTTAACCA AAATGGCGGG 120

TATTTATGCA TTCCCCGGAC AAACCCTGTG TATCGAGGGC CCTACTCGAA CCCCTACTCG 180

ACCCCTACT CAGGTCCGTA CCCAGCAGCT GCCCCACCAAC TCTCAGCTCC AACTATCCC 240

ACGATCTCCA GGCTCTTAT ATGCCGCTTT GGATACCAGA TGGATGAAAG CAACCAATGT 300

GTGGATGTGG ACGAGTGTGC AACAGATTCC CACCAAGTCA ACCCCACCCA GATCTGCATC 360

AATACTGAAG GCGGGTACAC CTGCTCCTGC ACCGACGGAT ATTGGCTTCT GGAAGGCCAG 420

TGCTTAGACA TTGATGAATG TCGCTATGGT TACTGCCAGC AGCTCTGTGC GAATGTTCCCT 480  
  
GGATCCTATT CTTGTACATG CAACCCTGGT TTTACCCTCA ATGAGGATGG AAGGTCTTGC 540  
  
CAAGATGTGA ACGAGTGTGC CACCGAGAAC CCCTGCGTGC AAACCTGCGT CAACACCTAC 600  
  
GGCTCTTCAT TCTGCCGCTG TGACCCAGGA TATGAACTTG AGGAAGATGG CGTCATTGC 660  
  
AGTGATATGG ACGAGTGCAG CTTCTTGAG TTCCTCTGCC AACATGAGTG TGTGAACCAG 720  
  
CCCGGCACAT ACTTCTGCTC CTGCCCTCCA GGCTACATCC TGCTGGATGA CAACCGAAGC 780  
  
TGCCAAGACA TCAACGAATG TGAGCACAGG AACCACACGT GCAACCTGCA GCAGACGTGC 840  
  
TACAATTAC AAGGGGGCTT CAAATGCATC GACCCCATCC GCTGTGAGGA GCCTTATCTG 900  
  
AGGATCAGTG ATAACCGCTG TATGTGTCCT GCTGAGAACCT CTGGCTGCAG AGACCAGCCC 960  
  
TTTACCATCT TGTACCGGGA CATGGACGTG GTGTCAGGAC GCTCCGTTCC CGCTGACATC 1020  
  
TTCCAAATGC AAGCCACGAC CCGCTACCCCT GGGGCCTATT ACATTTCCA GATCAAATCT 1080  
  
GGGAATGAGG GCAGAGAATT TTACATGCGG CAAACGGGCC CCATCAGTGC CACCCGGTG 1140  
  
ATGACACGCC CCATCAAAGG GCCCCGGGAA ATCCAGCTGG ACTTGGAAAT GATCACTGTC 1200  
  
AACACTGTCA TCAACTTCAG AGGCAGCTCC GTGATCCGAC TGCAGGATATA TGTGTCGCAG 1260

TACCCATTC

1269

Document Name: Abstract

**Abstract**

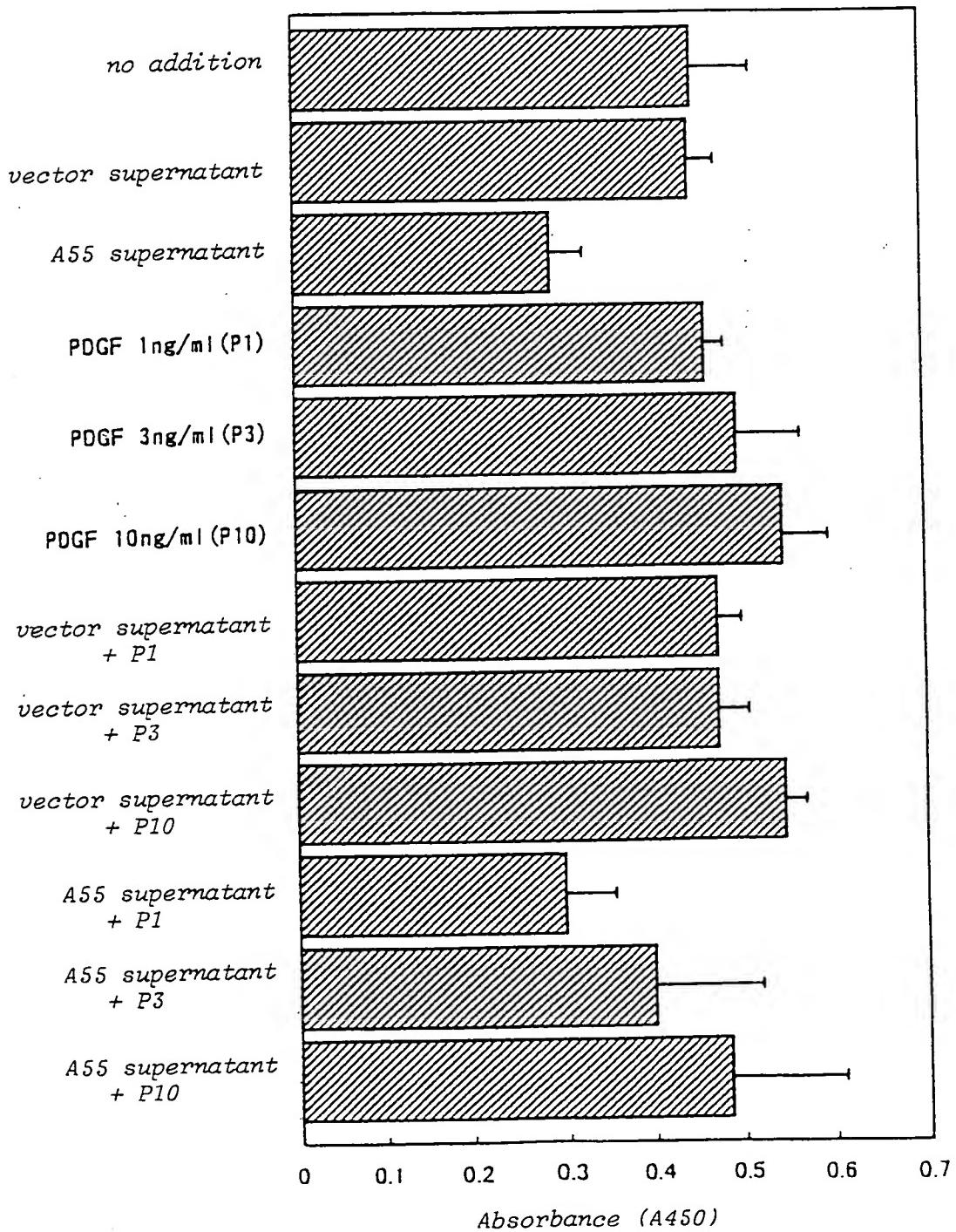
Constitution: Novel mouse and human polypeptide, a method for preparation of them, cDNAs encoding them, fragments capable to hybridizing them, a plasmid containing them, a host cell transformed with the plasmid, an antibody of the polypeptide, a pharmaceutical composition containing the polypeptide or antibody.

Effects: The polypeptide of the present invention possess an inhibitory activity on proliferation of vascular smooth muscle cells. So it is useful for the treatment of diseases induced by abnormal proliferation of smooth muscle, for example, arteriosclerosis, myosarcoma. It is also useful for the treatment and/or prevention of many kinds of diseases as the polypeptide have hematopoiesis regulating activity, tissue generation/regeneration activity, Activin/Inhibin activity, chemotactic/chemokinetic activity, hemostatic and thrombolytic activity, receptor/ligand activity etc.

Selected Fig.: none

Fig. 1

Inhibitory activity of mouse A55 protein on proliferation  
of rat aortic vascular smooth muscle cells stimulated by PDGF



09/674379

Fig. 2

Inhibitory activity of human A55 protein on proliferation  
of rat aortic vascular smooth muscle cells stimulated by PDGF

